STRIKING A BALANCE BETWEEN WATER FOR FOOD, ENERGY, AND THE ENVIRONMENT: A QUANTITATIVE FRAMEWORK TO GUIDE SUSTAINABLE WATER MANAGEMENT FOR A CHANGING FUTURE

BY

MARY A. YAEGER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Environmental Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor Murugesu Sivapalan, Chair and Director of Research
Professor Ximing Cai, Co-Chair
Professor Emeritus Gregory F. McIsaac
Yong Cao, Illinois Natural History Survey
Professor Atul Jain

Abstract

Humans and hydrology are strongly interconnected. Growing populations and improving technologies have led to a rising demand for fresh water, increasing the difficulty of finding a sustainable balance between competing water needs in a changing world. A major challenge to achieving this balance is the complexity of the interactions between two changing systems. While current research has provided insights into either the impacts humans have on hydrology or, conversely, hydrological variability on human society, the dynamic nature of these interactions is still not well understood. This dissertation develops and applies a framework for considering and analyzing the human and hydrologic systems as a linked, co-evolving system, utilizing insights developed from past interactions to guide exploration of future ones. The aim of this work is to further our understanding of these dynamic interactions between the human and hydrologic systems using a novel combination of modeling tools to investigate possible future behavior for patterns that reveal emergent properties.

This dissertation is presented around two main themes. The two studies of the first theme examine historical hydrologic data for patterns that provide insight into the underlying processes controlling catchment behavior. The first study comparatively analyzes many catchments to determine the dominant processes controlling the flow duration curve, while the second utilizes insights from the first to guide analysis of the historical co-evolution of two intensively-managed case study catchments to better understand human-hydrologic system interactions. Results from the first study suggested that the influence of catchment properties was strongest in the low flow tails of the flow duration curve, but also that in many places, including the case study region, the agricultural Midwestern US, humans were the dominant process. Building on this, the second study comparatively analyzed the two catchments as linked, co-evolving human-hydrologic systems and found their current hydrologic responses

to recent climate reflect their unique histories and are highly heterogeneous due to differences in their geologic pasts and human modifications such as artificial drainage and reservoir operation. These differences manifest temporally, from annual to daily time scales, and spatially, both within and between the watersheds, suggesting a complex, nonlinear response to other, future external driving forces.

The two studies of the second theme develop and then apply an integrated modeling framework to better address the dynamic interactions between the human and hydrologic systems. This framework employs a System of Systems optimization model to emulate human development decisions which are then incorporated into a watershed model to estimate the resulting hydrologic impacts. The two models are run interactively to simulate the co-evolution of coupled human-nature systems, such that reciprocal feedbacks between hydrologic processes and human decisions (i.e., human impacts on critical low flows and hydrologic impacts on human decisions on land and water use) can be assessed. The framework is applied to the two case study watersheds, in the context of proposed biofuels development. This operation is illustrated by projecting two possible future co-evolution trajectories, both of which use dedicated biofuel crops to reduce annual watershed nitrate export while meeting ethanol production targets. Imposition of a primary external driver (biofuel mandate) combined with different secondary drivers (water quality targets) results in highly nonlinear and multi-scale responses of both the human and hydrologic systems, including multiple tradeoffs, impacting the future co-evolution of the system in complex, heterogeneous ways. In one watershed, the strength of the hydrologic response is sensitive to the magnitude of the secondary driver; 45% nitrate reduction target leads to noticeable impacts at the outlet, while a 30% reduction leads to dominant impacts that are mainly local. The local responses are conditioned by previous human hydrologic modifications and their spatial relationship to the new biofuel development. This sensitivity is not as evident in the second watershed, where past human modifications to hydrology serve to both increase the importance of outlet flow and to partially mitigate some of the negative impacts of the external drivers. This emphasizes the importance of past co-evolutionary history in predicting future trajectories of change.

The work presented here provides some insights into the dynamic interactions that emerge from external driving forces (both climate and policy) propagating through connected human

and hydrologic systems, dependent on local geology and landscape properties and unique histories of anthropogenic modifications. This framework is a first step to a more integrated, dynamic approach to the study of coupled human and natural systems under change.

To my mom and dad, for their love and support.

To the memory of my uncle, Vernon Dyke, without whom I would not have been able to begin this journey.

Acknowledgments

I owe a debt of gratitude to many people without whose support along the way I would not be here, finishing this PhD.

First and foremost, I would like to give special thanks to my advisors, Murugesu Sivapalan and Ximing Cai, for their generous support of my work, their guidance towards my development as a researcher, and their patience when my progress was slow. I especially wish to thank them for providing me with unique, valuable research and academic opportunities along the way, and the freedom to pursue them.

I also wish to thank my mentors here for sharing their knowledge and providing support along the way: my committee members Gregory McIsaac, Yong Cao, and Atul Jain, for providing insightful comments and challenging questions; Praveen Kumar for asking the tough questions in group meetings and for informal discussions on research; and Gary Parker for all the jokes and for all the laughs when things looked bleak.

I also owe many thanks to my colleagues here in the lab: Ciaran Harman for being a great mentor and friend when I was still new and finding my academic way; Hongyi Li for his guidance and instruction; everyone I have collaborated with on papers and projects, especially Sheng Ye, Evan Coopersmith, Xiao Zhang, Ruijie Zeng, and Lei Cheng ("Charlie"); and my officemates Phong Le, Chuan Li, Jihua Wang and Debsunder Dutta for all those interesting discussions.

I would not have made it this far without dear friends outside of academia, especially Jennifer Staples and Michael Harber, who have been there since the beginning - I can't thank you enough; Steve McWhorter, who got me here, literally - and so much more; Joseph Hopper for his endless patience and for all the tech support with my beginner's programming questions; Miriam Cope and Christina Bouwens who got to the finish line ahead of me and

who've provided advice and much encouragement; and Brett and Beverly Tackitt, and my aunts Susan Dyke and Therese Ferguson for providing refuge and a place to recharge.

Lastly, I owe special thanks to my mother, Marie Yaeger, for her loving, unconditional support all through these years. Thank you, Mom.

Table of Contents

Chapter 1

Introduction

1.1 Research Motivation

As the global human population continues to grow and develop, the need for a stable, sustainable supply of clean water becomes ever more important. The technological advance of human societies brings with it a change in water needs; in general, development means a change from water for agriculture to water for both agriculture and energy production. With the global call for reduction in fossil-fuel derived CO_2 emissions, renewable energy sources such as biofuels are beginning to supplement more traditional water-for-energy uses, such as hydroelectric and nuclear power generation (King et al., 2008; Varghese, 2007). This increased water extraction and usage by agriculture for both food and energy often results in mounting pressure on, and potential degradation of, the environment (De Fraiture et al., 2008; Chiu et al., 2009; Gerbens-Leenes et al., 2009). Even in the absence of competition from bioenergy, conflicts between ecosystem water needs and those of agriculture can be severe; on the Klamath river in the northwestern USA (Levy, 2003) and the Murrumbidgee river in Australia (Connell and Grafton, 2011; Sivapalan et al., 2012) when environmental water requirements were finally defined and enforced, much of the allocation for agriculture was diverted, resulting in economic losses from the agricultural sector.

In order to manage these finite water resources in a sustainable manner, it is imperative that we find some balance between the changing water needs of society and ecosystems (Wallace et al., 2003). The importance of this is further underscored by the uncertainty of climate change; the past (in a statistical sense) may no longer be a good predictor of the future (Milly et al., 2008). Therefore, deterministic modeling methods based on this stationarity assumption may not provide realistic predictions of the future state of the system,

especially over long time scales (Peel and Blöschl, 2011). Thus there is a growing need for a new framework within which to mediate this competition between human and environmental water needs (Wagener et al., 2010; Sivapalan et al., 2011a). This problem of balancing the water needs of agriculture, bioenergy production, and the environment in a sustainable fashion under uncertain conditions provides the motivation for this research, the goal of which is to better understand how the human and the environmental systems coexist and interact.

1.2 Methodology framework

In this study, we first analyze the data and historical record in a comparative fashion to better understand the processes and interactions that have already occurred. Previous studies have identified many interactions between the human and natural systems, but the underlying processes that govern them may still be undiscovered (Liu et al., 2007b; Wagener et al., 2010). Recently, there has been recognition of the need for understanding these interactions within coupled systems (DeFries and Eshleman, 2004) and a focus on frameworks whose aim is discovery and understanding of the underlying processes (Sivapalan, 2006; Sivapalan et al., 2012; Thompson et al., 2013; Harman and Troch, 2014). A number of studies have accomplished this through comparative analysis of data to identify patterns, which are the emergent outcomes of these processes acting over time (Troch et al., 2009; Basu et al., 2010; Cheng et al., 2012; Coopersmith et al., 2012). Investigation and analysis of these patterns have led to a better understanding of the underlying processes (Basu et al., 2011; Sivapalan et al., 2011b; Ye et al., 2012; Yaeger et al., 2012). Furthermore, the insights gained from the large scale comparative studies, e.g. Yaeger et al. (2012), can provide guidance to more focused, smaller scale studies, e.g. Yaeger et al. (2013), in forming process-based understanding of the interactions between systems. Therefore, the comparative hydrology approach provides a quantitative foundation for this study.

Because of the complexities of the problem of sustainably meeting competing water needs in an uncertain future, human and environmental systems need to be considered and studied as interconnected parts of a whole. Liu et al. (2007b) emphasize this point by outlining three aspects that must be integrated in the research of coupled human and natural systems

(CHANS):

- Discovery of processes that link the human and environmental systems
- Inclusion of bi-directional interactions and feedbacks between humans and the natural system
- Understanding of how these interactions and feedbacks lead to emergent properties unique to the coupled system.

Thus in order to capture the underlying dynamics of CHANS, the interactions and feedbacks of both systems need to be incorporated in any modeling exercise, the focus of which should be on the co-evolution of the system, both in the past and projected into the future, so that possible unexpected patterns or behavior can be detected (Sivapalan et al., 2012). Improved understanding and predictive insights developed from this approach can provide guidance to water resource management decisions under uncertain, changing conditions (Thompson et al., 2013). Because this study seeks to develop such predictive insights, we incorporate these principles of CHANS into the integrated modeling framework developed herein.

1.3 Research context: Biofuels case study

This research is a small part of a much larger project that seeks to develop strategies to sustainably operate and expand the interdependent infrastructure systems of the emerging bio-economy, including water and energy supply, transportation, and focusing on the production of bioethanol from cellulosic feedstocks (Figure 1.1). The goal of this larger project is to integrate the many subsystems into an overarching optimization model in order to analyze the interdependence and sustainability of biofuels engineering infrastructures, thereby gaining deeper understanding of the dynamics of this coupled system. The present study incorporates this multisystem optimization model into a CHANS framework to provide a closer look at the interactions between the human and hydrologic systems.

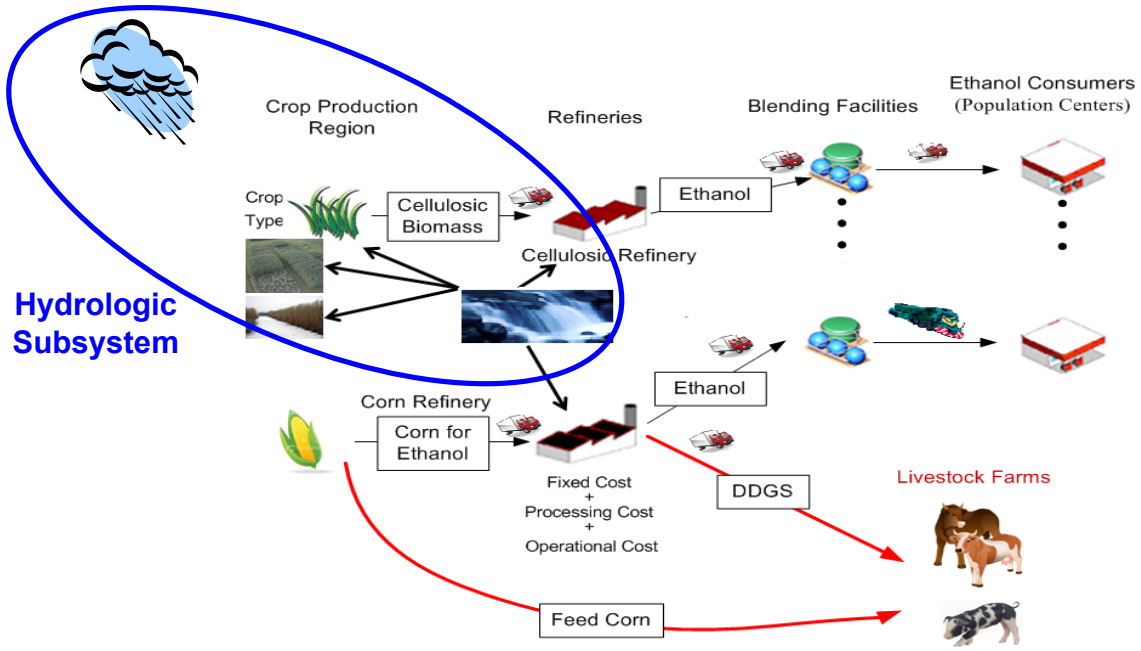

Figure 1.1: Conceptual model of the interconnected subsystems of the biofuels system: environment, farm, transportation, refinery, economy.

In order to determine the impacts and feedbacks within a CHANS framework, an assessment must be made of the states of both the human and environmental systems. In the context of the biofuels case study, it is expected that the low flows be most affected by the proposed biofuels development. Since low flow periods can be a time of aquatic stress, especially if they occur in summer (Schlosser, 1991), a minimum threshold flow has been a commonly-used method for determining the condition of the river ecosystem (Richter et al., 1997). In some cases, this may be an inadequate indicator and as a result the human system might miss potentially crucial ecosystem feedbacks. By considering the ecosystem as well as the hydrology in the environmental response, different feedbacks might emerge, furthering understanding of the interactions between human and natural systems, which would enable a more informed human response. While this study utilizes a threshold minimum flow to demonstrate the framework, additional work towards incorporating ecological needs into the assessment of the natural system is also included.

1.4 Research Objectives

The aim of this research is to develop a generalized integrated modeling framework for coupled human and nature systems that specifically incorporates the bi-directional interactions and feedbacks characteristic of these systems. We then apply this framework to the case study CHANS to address specific research questions or aspects of the problem.

- What insights can be extracted from the history of change in the watershed as expressed in the present-day catchment response, regarding possible responses to future, unknown changes?

In this study we identify emergent patterns that are the signatures of past changes and use these insights to identify feedback loops of change, impact, and response in the historical record of a smaller-scale coupled agricultural-hydrologic system.

- By incorporating the feedbacks from the environmental system onto the human system and the response of the human system to the environment, what does the behavior of the coupled system through time reveal about the underlying dynamics of the system?

We employ the previous findings in the development of an interactively-run integrated modeling framework that combines a systems optimization model with a semi- processed-based hydrologic model and impose an external driving force (here, a mandate for biofuels development) that is then propagated internally through the linked systems such that reciprocal feedbacks between hydrologic processes and human decisions (i.e., human impacts on critical low flows and hydrologic impacts on human decisions on land and water use) can be assessed. These feedback loops are projected forward in trajectories of change to investigate and quantify watershed and human system responses through time.

- How to use ecogeomorphology based on some simple geomorphologic parameters to translate streamflow from a hydrologic model into ecological habitat indicators?

In keeping with the goal of process-based understanding, we use geomorphologic data and fish sampling data to look for a correlation between hydraulic geometry and fish ecological

indices in order to incorporate both ecology and hydrology in the quantitative assessment of the state of the environmental system in the framework.

The original contributions from this research pertain to the conceptualization of the multisystem interactions and feedbacks in the case study CHANS through development and application of a quantitative integrated modeling framework, as well as the predictive insights gained regarding the processes linking these systems.

1.5 Thesis outline

This dissertation is structured as a series of journal manuscripts, where each chapter is a separate manuscript that has either been published, submitted for publication, or is in preparation for submission to a journal. The chapters are organized along two main themes, comparative analysis of data and the development and application of the modeling framework, where the first serves to inform the second. Chapter 2 synthesizes the findings of a large-scale comparative hydrology study encompassing 200 catchments across the US, focusing on the flow duration curve (FDC) and in particular the importance of the low flow tails in determining catchment signatures of change. In Chapter 3, the comparative hydrology approach is applied to the biofuels case study watersheds to identify signatures of change in their historical trajectories. Chapter 4 introduces the integrated modeling framework and applies it to one of the case study watersheds. In Chapter 5, the integrated modeling framework is applied to both case study watersheds and the trajectories of change are comparatively analyzed. Chapter 6 presents general conclusions and future research directions arising from this study. Appendix A contains the results of the regression analysis of the ecological and geomorphological data.

Chapter 2

Exploring the physical controls of regional patterns of Flow Duration Curves: A synthesis of empirical analysis, process modeling and catchment classification

Abstract

This paper reports on a four-pronged study of the physical controls on regional patterns of the Flow Duration Curve (FDC). This involved a comparative analysis of long-term continuous data from nearly 200 catchments around the US, encompassing a wide range of climates, geology, and ecology. The analysis was done from three different perspectives statistical analysis, process-based modeling, and data-based classification, followed by a synthesis, which is the focus of this paper. Streamflow data was separated into fast and slow flow responses, and associated signatures, and both total flow and its components were analyzed to generate patterns. Regional patterns emerged in all aspects of the study. The mixed gamma distribution described well the shape of the FDC; regression analysis indicated that certain climate and catchment properties were first order controls on the shape of the FDC. In order to understand the spatial patterns revealed by the statistical study, and guided by the hypothesis that the middle portion of the FDC is a function of the regime curve (RC, mean within year variation of flow), we set out to classify these catchments, both empirically and through process-based modeling, in terms of their regime behavior. The classification analysis showed that climate seasonality and aridity, either directly (empirical classes) or through phenology (vegetation processes), were the dominant controls on the RC. Quantita-

This work has been published as: Yaeger, M., Coopersmith, E., Ye, S., Cheng, L., Viglione, A., and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves – Part 4: A synthesis of empirical analysis, process modeling and catchment classification, Hydrol. Earth Syst. Sci., 16, 44834498, doi:10.5194/hess-16-4483-2012, 2012. All figures, tables and data were created by Mary Yaeger unless otherwise indicated.

tive synthesis of these results determined that these classes were indeed related to the FDC through its slope and related statistical parameters. Qualitative synthesis revealed much diversity in the shapes of the FDCs even within each climate-based homogeneous class, especially in the low-flow tails, suggesting that catchment properties may have become the dominant controls. Thus, while the middle portion of the FDC contains the average response of the catchment, and is mainly controlled by climate, the tails of the FDC, notably the low-flow tails, are mainly controlled by catchment properties such as geology and soils. The regime behavior explains only part of the FDC; to gain a deeper understanding of the physical controls on the FDC, these extremes must be analyzed as well. Thus, to completely separate the climate controls from the catchment controls, the roles of catchment properties such as soils, geology, topography etc., must be explored in detail.

2.1 Introduction

Catchment signatures quantify hydrologic responses to rainfall inputs in a compact manner; by distilling catchment behavior into a few signatures, classification of variable behavior across many different catchments can be made. One such signature, the regime curve (RC), describes the intra-annual variability of monthly (or even daily) average (ensemble mean) streamflows. Another signature, the flow duration curve (FDC), plots daily streamflow magnitude (on a log scale) as a function of the percent of time it is exceeded. Encoded within these signatures are the combined impacts of climate, geology, topography, ecology, and even human activities. We hypothesize that some function relates each hydrologic signature, the FDC included, to both the climate and the landscape, thus connecting the variability of climate inputs to the variability of runoff. This function can then be determined empirically, through the application of complex process-based models, or some combination thereof, both in detail for a small region via a process-based (Newtonian) approach, or through a comparative, data-based (Darwinian) approach. On the basis of these analyses, unique functions can then be defined for regions of similar catchment and landscape properties. With this idea of regionalization based on clusters of similar behavior, the prediction of the FDC or any other hydrologic signature can be greatly improved in cases where this function

is not easily determined, such as ungauged basins.

The FDC is a signature of the within-year variability of the runoff in a watershed and represents the relationship between magnitude and frequency of streamflows at some location in the basin (Botter et al., 2008). Although the FDC is widely used in practice, and with much empirical knowledge in specific regions having been gained, the physical controls on the shape of the FDC are still not well understood (Yokoo and Sivapalan, 2011). Because of its wide usage and the simple data demands on its construction, the FDC is an important catchment signature. In a rapidly-changing world, such signatures can be used to track catchment responses to perturbations, in addition to the more traditional application to analyze or interpret long-term catchment responses (Vogel and Fennessey, 1995). Improved understanding of the physical processes controlling the shape of the FDC would allow for better comparisons of catchments across diverse regions as well as over long time periods for a single catchment. Ultimately, an increased understanding of the controls on the FDC would enable better estimation and prediction of the response from ungauged catchments, as well as improving our ability to carry out hydrologic space-for-time substitutions when making predictions under change.

In the past, FDCs have been generally analyzed through purely graphical means, e.g., Vogel and Fennessey (1994, 1995) or stochastically by fitting appropriate statistical distributions e.g., LeBoutillier and Waylen (1993); Castellarin et al. (2004, 2007). Recently, the stochastic approach has been augmented with a more process-based approach, wherein the different components of the catchment's dynamic response are incorporated into stochastic models (Botter et al., 2007a,b, 2009; Muneepeerakul et al., 2010), thus allowing for a more complete picture of the streamflow variability of different catchments. Their complete model contains many underlying physical properties of a catchment but still uses a Poisson assumption for the arrival of rainfall events, which enables an analytical solution to be obtained. The model is applied seasonally, and since there is no carryover of soil moisture storage between seasons, the results are not applicable to the whole year. This may not be realistic in all cases, especially in regions that exhibit strong seasonal variations of climate and water storage, and highlights the need for a more general framework, one for the entire year that captures within-year variations in climate and soil moisture storage.

The numerical study of Yokoo and Sivapalan (2011) followed this process-based framework with one important exception: there is no Poisson assumption. Rather, their numerical water balance modeling framework uses precipitation data (either observed or synthetic) to both provide the randomness inherent in precipitation at the event scale and yet capture the seasonal variability of climate. Carryover of soil moisture storage across the year allows for antecedent soil moisture conditions to affect the runoff response of the catchment. As a result of this work, they hypothesized that the flow duration curve could be divided into three parts, or limbs. The upper limb corresponds to the fast flow response of a catchment to precipitation; the middle limb is related to the slow flow response of the catchment, and as such is a function of the monthly regime curve; and the tail limb could be related to evaporation from variable saturated regions in the catchment and impacting low flows. This insight, to separate the streamflow into fast flow and slow flow and look at the FDC the same way, provided the motivation for this four-pronged study, which was undertaken to further explore the possible physical controls on the FDC. While the Yokoo and Sivapalan work was mainly hypothetical and used synthetic inputs to a hydrological model, the present four-pronged study that follows it is empirically based, using long-term continuous daily climate and streamflow data from about 200 catchments in the MOPEX database. Using this data, the possible controls on the FDC were explored from three angles: statistically to identify patterns, through the use of a simple physically-based model to separate climate and landscape controls, and through catchment classification to organize the patterns into classes.

The first angle of exploration was empirical (Cheng et al., 2012), and was accomplished by first separating total streamflow (Q) into fast flow (Q_f) and slow flow (Q_u) and fitting each of their duration curves as well as that of precipitation (P) to a mixed gamma distribution, which has three parameters: κ, θ, and α. The first two parameters - the shape parameter κ, which affects the shape of the curve, and the scale parameter θ, which stretches or shrinks the curve and is related to the magnitude of the variable in question are embedded in the standard gamma distribution. However, for the duration curves of P and Q_f, as well as (Q of catchments in arid regions, the duration can be much less than 100%. Thus a third parameter, α, the fraction (or ratio) of no rain or flow days, was introduced to capture

this (see Cheng et al. (2012) for further details of the fitting). Spatial plots of the gamma parameters obtained for the long-term (covering the entire record) duration curves for all catchments as well as the annual duration curves of a few selected catchments showed interesting regional (between catchment) patterns, and between-year variations. Moreover, both the between-catchment and between-year analyses demonstrated that the physical properties of the catchment, represented in the baseflow index (BI), are the dominant controls on the shape of the FDCs of Q and Q_u, while climate (precipitation) is the dominant control on the shape of the FDC of Q_f. However, this dominance was by no means absolute, which meant that other factors may be controlling the shapes as well, and this in turn led to the modeling portion of the study as a way to further elucidate these other controlling factors.

The second angle of exploration involved top-down process modeling (Ye et al., 2012), using a simple two-stage hydrologic model following the modeling framework outlined in Yokoo and Sivapalan (2011). The regime curve (RC) is used as the signature of comparison during the model calibration, since as previously mentioned, the middle portion of the FDC, which connects the high flows to the low flows, is related to the RC, and both the FDC and the RC are necessary for understanding the catchment response. Where the base model failed to capture fully the shape of the RC, processes were added to the model in a systematic fashion, but always guided by the data. Controlling processes were then identified for each catchment, and when plotted spatially, these also showed interesting regional patterns. As an example, phenology (vegetation-induced seasonality in streamflow) and snowmelt processes dominated the Northeast portion of the US, with both becoming less dominant towards the south and west. In some cases, however, very different catchments had the same controlling process. For instance, precipitation seasonality dominated in both the mountainous Pacific Northwest and in the flat peninsula region of Florida, highlighting the fact that dominant processes alone may not make catchments similar, and pointing towards a need for further classification of catchment response.

The third angle of exploration involved the development of a classification system using information readily obtained from the daily RC (Coopersmith et al., 2012). Unlike the previous two studies, all 428 catchments in the MOPEX dataset were used in order have the highest possible variation in climate and location, since most of the long-term hydrologic

data in the US tends to be concentrated in the more humid, eastern half of the country. As with the modeling study, the RC was chosen as the signature of comparison instead of the FDC and for much the same reasons, with one addition: the RC contains much more easily interpretable information than does the FDC. This may indeed be one of the reasons the FDC is still not well-understood, even after much use and much study. From the RCs of P, potential evaporation (E_p), Q and its components Q_f and Q_u, one can establish the average aridity and seasonality of the climate, as well as when P and Q peak for a particular catchment. Using these four similarity indices and an objective function that minimized within-group variance, catchments were clustered along a decision tree until further splitting ceased when catchments within a terminal node were sufficiently similar or there was no further reduction in variability by splitting again. Over 3/4 of the catchments studied fell into only six classes, with the first major distinction between catchments being seasonality. The eastern portion of the US experiences very low seasonality in precipitation, with a wet/dry seasonal pattern becoming more pronounced as one moves further west across the continent. Again, interesting regional patterns emerged when the catchment classes were displayed spatially. As with the model processes, the Northeastern US formed a different cluster than the Southeast and the Midwest US formed two major clusters, but this time, the Pacific Northwest and Florida in the extreme Southeast fell into two different classes. This shows that while both regions experience high seasonality in P (the dominant model process), the amount of rainfall each region receives (more overall in the NW) as well as the timing of the wet season (in the NW it is out of phase with E_p, while FL it is in phase) distinguish Florida from the Pacific NW. This analysis leads to the conclusion that there may be some connection between the results of each of these studies, and elucidation of this connection could lead to deeper insights into the controls on the FDC.

Each approach outlined above produced valuable insights into the controls on the FDC and showed tantalizing spatial patterns that pointed toward some underlying structure or controlling process(es). However, since the main focus of the modeling study and the classification scheme was the RC, not the FDC, there still remains a need to extract information from these three studies and apply it to the FDC and its possible controls, i.e. to synthesize the results. What, then, is synthesis? Because catchments are complex systems that have

co-evolved over time to reach their present state, the problem of predicting future behavior can quickly become multi- dimensional. Since a catchment can be thought of as a "filter" of the precipitation inputs, with various interactions and feedbacks among the various processes at work within the catchment, one way to deal with this problem of complexity is to look for emergent patterns in the catchment signatures, and thereby gain some fundamental understanding of catchment functioning. Sivapalan et al. (2011a) in their call for a synthesis approach to hydrology define these patterns as consistent trends of similarity or difference across time and/or space. A focus on patterns in order to gain understanding of the underlying processes and a bringing together of various research efforts into one integrated body of work forms the framework for our synthesis approach. Using the hypothesis of Yokoo and Sivapalan (2011) as the basis for analysis and combining the results from the first three parts of our study of hydrologic signatures, the synthesis approach is applied to the current problem. The goal of this paper, the last in the four-part series outlined in the preceding paragraphs, is to map the possible connections between the parameters of the mixed gamma distribution obtained in the empirical study, the dominant model processes, and the catchment classification. It is in finding the common ground between the empirical, modeling, and classification work, and building further on this synthesis that a deeper understanding of the physical processes controlling the shape of the FDC can be gained.

This chapter is divided into four sections. In Section 2.2, we have introduced the FDC, and the ideas of regionalization, similarity, and synthesis as a way to better understand the physical controls on the shape of the FDC. Section 2.3 describes the study site, data used, and the specific methods used in the synthesis, and Section 2.4 presents the results of the synthesis. In Section 2.5 the conclusions of the entire four-part study are summarized. We conclude with a discussion of the limitations of the study and possible future work that will overcome these limitations.

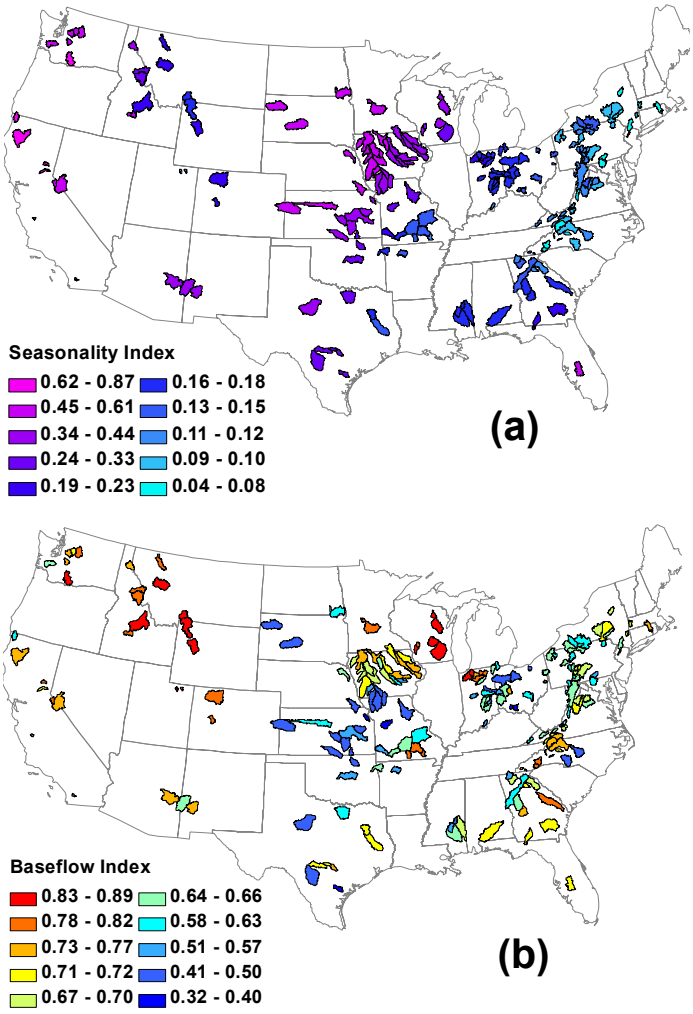

Figure 2.1: Spatial patterns of seasonality index (a) and baseflow index (b) across the continental US.

2.2 Data and Methods

2.2.1 Study Site and Data Description

Out of the 428 catchments in the MOPEX database, 197 catchments were chosen for their long-term, continuous daily data sets. The data used in the present work (area-averaged daily precipitation, total streamflow, baseflow-separated streamflow, and potential evaporation, all in units of mm) was assembled into a processed database as part of the Hydrologic Synthesis Project (Sivapalan et al., 2011b; Brooks et al., 2011) and is freely available at

http://voda.hwr.arizona.edu/mopex/. While the overall spatial distribution of these catchments covers the entire continental United States and therefore represents most of the climate and geographic variability found there, a majority of the catchments are located east of the Mississippi River (Figure 2.1). In the empirical, modeling, and classification studies, the common theme of emergent regional patterns in the spatial distribution of the gamma parameters, the dominant model processes, and the catchment classes provided visual evidence of some underlying organization. Clear spatial patterns are also seen in climate indicators such as the aridity index (E_p/P) which exhibits the typical continental pattern of increasing aridity with increasing distance from the coast and with decreasing latitude in the Western US, as well as the seasonality index (Fig. 2.1a), adapted from Walsh and Lawler (1981) and defined here as:

$$SI = \frac{\sum_{i=1}^{365} |P_i - \sum_{i=1}^{365} P_i|}{\sum_{i=1}^{365} P_i} \tag{2.1}$$

where P_i is the daily precipitation at each site. Less clear is the spatial organization of catchment-related indices such as the baseflow index , shown in 2.1b. Since the baseflow index is related in large part to geology, which tends to vary at the local scale, its spatial pattern shows a more localized clustering compared to the climate indices, which tend to vary along a gradient.

Daily data from the processed MOPEX database was first normalized before being used in this study, with P normalized by the mean annual daily precipitation for the 54 year time period, and Q, Q_f, and Q_u all normalized by the corresponding mean flow values for the same time period. The duration curves of P, Q, Q_f, Q_u analyzed herein were produced using the quantile method (Vogel and Fennessey, 1994), i.e., only data points corresponding to the 0.5, 1.0, 1.5, , 99.0, and 99.5 exceedance percentages from the 54 years of normalized daily data were considered. In addition to the processed daily time-series data, results from the first three parts of the four-part study described previously were used to create a database for the synthesis analysis. From the empirical analysis (Cheng et al., 2012), the fitted parameters of the mixed gamma distribution α, κ, and θ for P, Q, Q_f, and Q_u for each catchment, were obtained. From the modeling study (Ye et al., 2012), the dominant model process

"class" (i.e., that combination of processes which were found to be necessary for good model prediction for a given catchment) for each of the 197 catchments was extracted. From the catchment classification study (Coopersmith et al., 2012), the class associated with each of these catchments was added to this new database. Lastly, daily regime curves used in the previous two studies will also be used. These are calculated using the entire period of record by finding the average flow for each Julian day of the year, then smoothing with a 30-day circular moving average (see Coopersmith et al. (2012) for further details).

2.2.2 Synthesis Methods

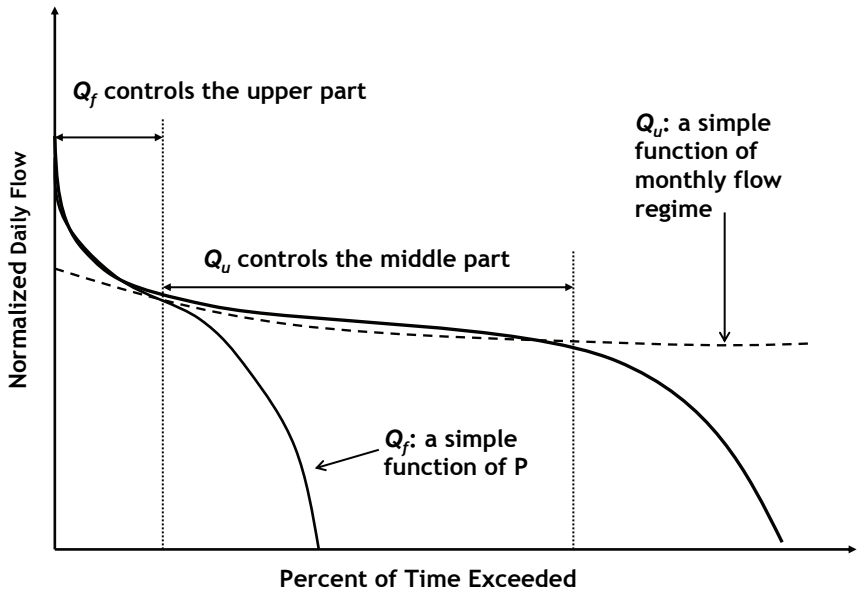

Figure 2.2: Conceptual model of the controls on the FDC (adapted from Yokoo and Sivapalan (2011)); here Q_f is the fast flow (surface runoff) and Q_u is the slow flow (sub-surface runoff).

As discussed earlier, all the work done in these related studies, including the present one, has been guided by the framework laid out in Yokoo and Sivapalan (2011), which is presented visually in Figure 2.2. This conceptual model was built on the partitioning idea outlined by L'vovich (1979), wherein P is partitioned into Q_f and catchment wetting (infiltration),

and this wetting is then further partitioned into Q_u and vaporization (ET). Thus the upper portion of the FDC, which represents the high flows that are exceeded rarely, is controlled by Q_f, a simple function of P. The middle portion of the FDC, which represents the average flows, is controlled by Q_u, which is a function of the monthly flow regime. The lower portion of the FDC, which represents the extreme low flows that are exceeded most of the time, may be a function of evaporation from saturated areas, among other things. In this framework, the RC, which contains the average flows experienced by a catchment throughout the year, forms the bridge between the low flows and the high flows in the FDC. Thus the results from the modeling and classification studies can be used to learn more about the physical controls on the FDC.

Because the middle limb of the FDC forms the connection between the extreme high flows and the extreme low flows, and is related to the average flow regime of a catchment, the slope of this portion of the FDC is often used in modeling and classification studies as a means of quantifying the shape of the FDC (Carrillo et al., 2011; Sawicz et al., 2011; Yadav et al., 2007). The fitted gamma parameters also quantify the shape of the FDC, but they encompass all of the FDC, not just the middle portion. Since the middle part of the FDC relates to the RC and therefore the modeling and classification studies, it is included here as a way to synthesize the findings of these studies. From the three papers cited above, the slope of the flow duration curve is defined as:

$$Slope_{FDC} = \frac{\ln(Q_{0.33}) - \ln(Q_{0.66})}{(0.66 - 0.33)} \tag{2.2}$$

where $Q_{0.33}$ and $Q_{0.66}$ are the flow values exceeded 33% and 66% of the time, respectively. Since P and Q_f are either zero or some extremely small number (<0.001 mm) and usually exceeded at or less than 66% of the time in nearly all the catchments, the slopes (as defined here) of these FDCs were not mathematically meaningful, and thus only the slopes of Q and Q_u were used for this analysis.

The three previous studies - empirical, modeling, and classification - form three angles of what can conceptually be described as a synthesis "triangle" (Figure 2.3). The parameters of the mixed gamma distribution form the apex of the triangle, and at the base, moving

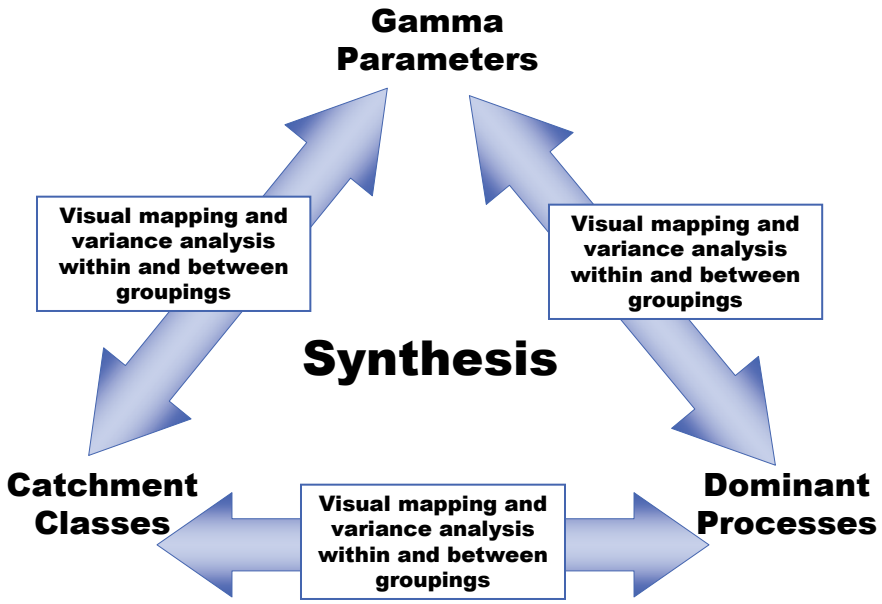

Figure 2.3: Connections to be explored between the three aspects of the FDC study.

clockwise from the top, are the dominant model processes and the catchment classes. The goal of the synthesis analysis is to make connections between each of these three angles, thus forming the "legs" and completing the triangle. To map the FDC to possible controlling processes, a two-stage analysis is conducted: for each leg of the synthesis triangle, a qualitative visual grouping analysis is followed by a quantitative variability analysis. Additionally, because two of the legs of the triangle are derived from the RC, the synthesis analysis will zoom in on the middle portion of the FDC and the same visual grouping and variability analysis will be performed on the slope of the FDCs sorted by both catchment class and dominant processes. Lastly, because of the regression relationship seen between the shape parameter of the mixed gamma distribution κ and the baseflow and seasonality indices, the slope of the FDC will also be compared to these indices. Since κ represents the entire FDC and the FDC slope represents just the middle portion of the FDC, a strong relationship may help clarify the underlying controls. Thus the analysis of the slope of the FDC builds on the three-way mapping outlined by the synthesis triangle.

The variability analysis must be done between sets of quantitative data (parameters of the

mixed gamma distribution, the slopes and quantiles of the FDCs) as well as between sets of qualitative data (the dominant model processes and catchment classes). For the quantitative data, this is accomplished by determining self-organizing behavior with a decision tree (Breiman et al., 1993). Consider a dataset of size n, with variance σ^2. Next, partition that dataset into m subsets, using an existing decision tree. Each subset i, of size S_i is characterized by a variance σ_i^2. Let a variance measure of the data, using the partitions implied by the tree, be calculated as follows:

$$\sigma^2 = \frac{\sum_{i=1}^{m} S_i \sigma_i^2}{n} \tag{2.3}$$

In essence, the average of the variance of each subset is calculated, weighted by the number of members of each subset. For the qualitative data, this is accomplished by determining self-organization into distinct classes. Again, consider a dataset of size n, which now is characterized by k distinct, qualitatively described classes. Let the entropy of any group of that data be defined as:

$$E = -\sum_{j=1}^{k} p_j \log_k p_j \tag{2.4}$$

where p_j represents the proportion of the group of class j. Equation (2.4) returns a value between 0 (minimum entropy, all elements within a set of the same class) and 1 (maximum entropy, equal proportions of a set of each of the k classes). Next, partition the dataset of size n into m subsets of size S_i, each with an entropy E_i, calculated with Eq. (2.4). Let the entropy of the data, using the partitions implied by the tree, analogous to Eq. (2.3), be calculated as follows:

$$E = \frac{\sum_{i=1}^{m} S_i E_i}{n} \tag{2.5}$$

If the variance measure calculated in Eq. (2.3) and the entropy calculated in Eq. (2.5) are lower than that of the entire dataset, we can conclude that the tree structure explains, in part, the variability of the parameters/classes we desire to map.

2.3 Results: Three-way Mapping and Synthesis

2.3.1 Catchment Classes and Dominant Processes

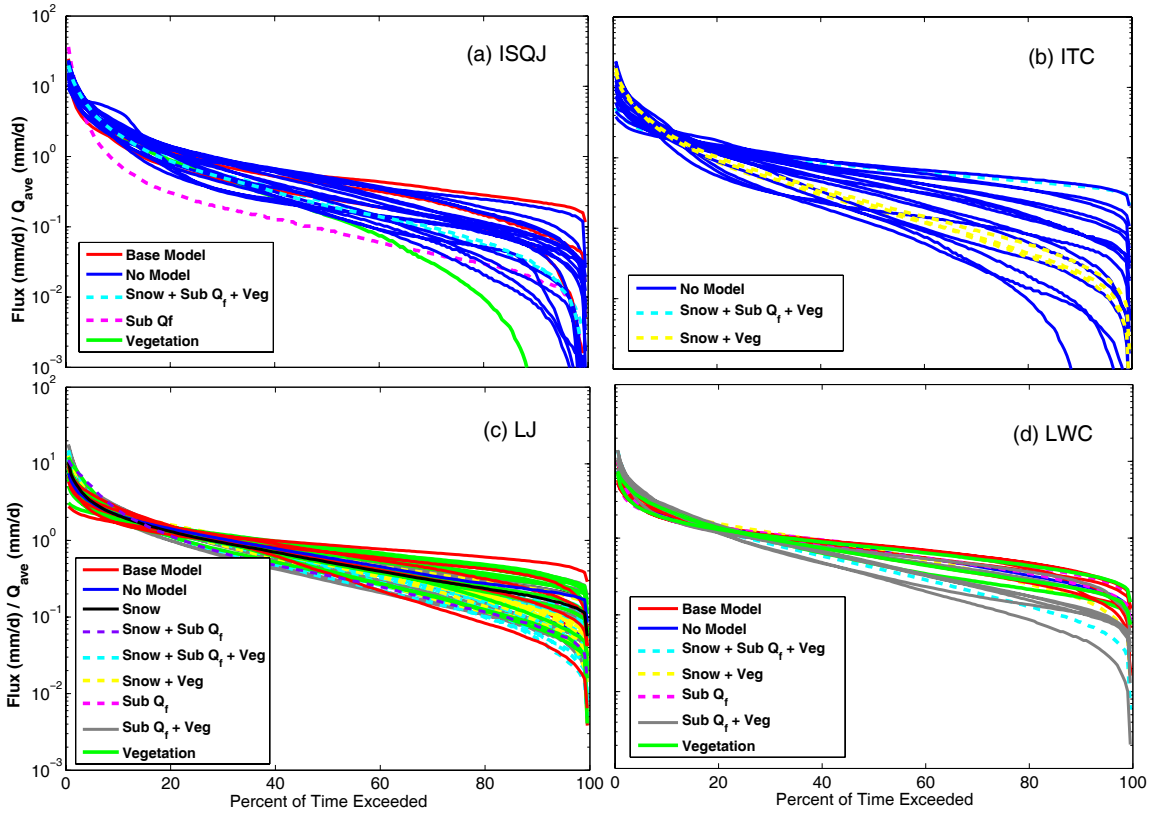

Figure 2.4: FDCs for total streamflow within the 4 main catchment classes, grouped by model process class (Note that (a) and (b) represent drier catchments with mildly seasonal rainfall P, while (c) and (d) represent wetter catchments with little seasonality in P).

Figure 2.4 shows the FDCs of catchments in the four largest catchment classes, which together account for about 2/3 of all the catchments in the study, grouped within each class by dominant model process. The class LJ (low seasonality, wetter spring), which includes 68 catchments and encompasses most of the Atlantic coast region of the US and the class LWC (low seasonality, wetter winters) which covers the southern Appalachian Mountain region and the southern Gulf states show remarkably little difference between FDCs for high flows exceeded about 15 - 20% of the time (Figure 2.4, c and d). The curves begin to diverge

at the mid- to low-flow end, as well as at the extreme high flow end, but even there the range is less than that of the intermediate seasonality classes ISQJ and ITC which include catchments in the Midwest from the humid corn belt region in Illinois and Iowa to the drier Plains regions south and west of the corn belt (Figure 2.4, a and b). Of note is that the catchments in these two intermediate seasonality classes are also where the simple process model had difficulty in satisfactorily reproducing the regime curve (Ye et al., 2012), while the two dominant low-seasonality classes are largely made up of model process classes relating to forest and snow processes.

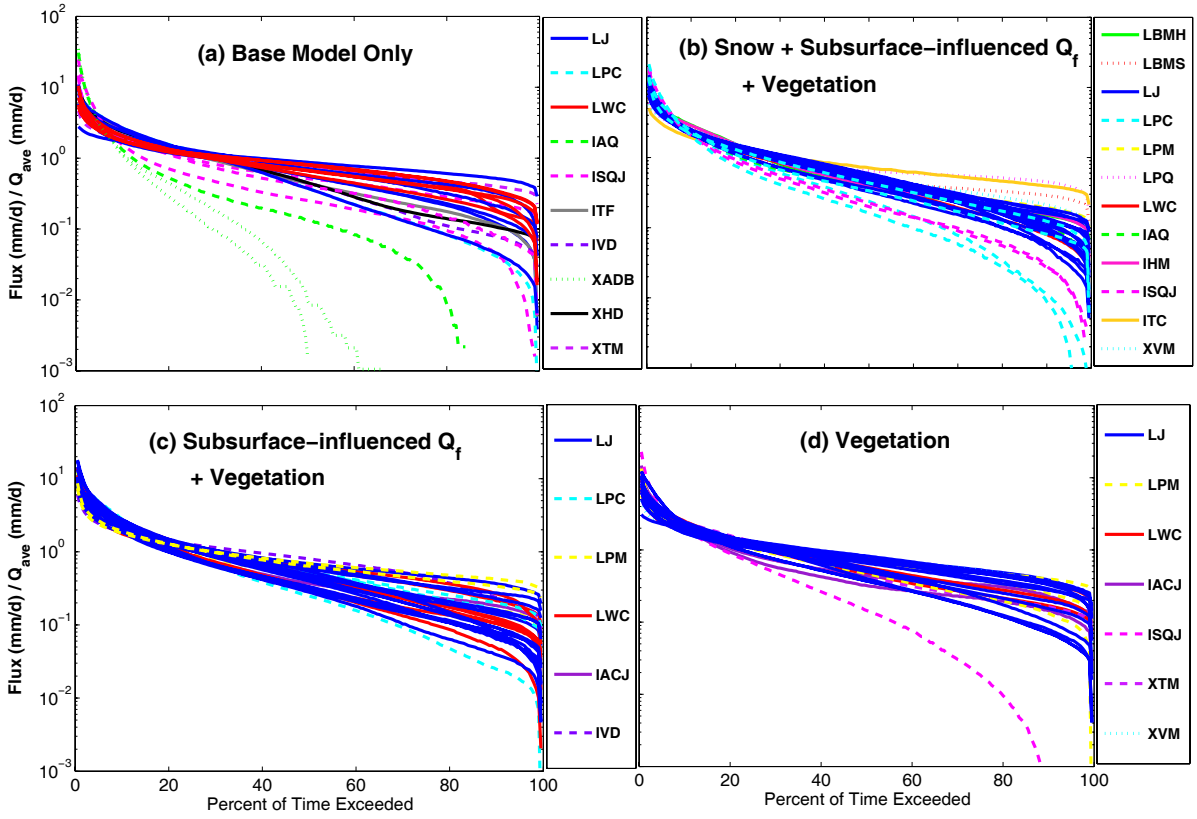

Figure 2.5: FDCs for total streamflow within the 4 main dominant model process classes, grouped by catchment class. (L, I, and X denote low to high seasonality in P, while V, H, T, S, and A denote low to high aridity).

In Figure 2.5 the grouping is reversed; here the FDCs of catchments in the four largest dominant model process classes, which together account for about half of all the catchments in the study, are grouped within by catchment classification. Again, we see two of these classes (b and c) have smaller range of FDC variability than the other two. A closer inspec-

tion reveals that both (b) and (c) seem to taper at the top (largest flows) while (a) and (d) do not. Since the vertical axis of an FDC is on a log scale, any differences that appear small at this end are actually quite large, while differences between FDCs at the low end, which appear large, are actually very small. It can be further noted that both (b) and (c) contain subsurface-influenced Q_f, and it may be due in part to this underlying process control that (b) and (c) have more similar high flows. Also of note in Figure 2.5 is that three of the four main dominant model processes involve vegetation (interception, phenology, or both) and thus the two large classes LJ and LWC, which are dominated by catchments in the NE and mid-Atlantic regions of the US, feature prominently in these groupings (Figure 2.5, b, c, and d). The base model cluster features the majority of the extreme seasonality classes as well as the more seasonal of the intermediate classes. In these catchments, the strong seasonal signal in precipitation overrides even the vegetation signature in the streamflow regime curve; the effect on the FDC is not so evident. However a lack of precipitation shows up very strongly in the FDC, where the drier catchments in the southern and western regions (e.g. IAQ and XADB in 2.5a, LPC and ISQJ in 2.5b, and ISQJ in Figure 2.5d) have a markedly different curve than their more humid counterparts.

The visual analysis of Figures 2.4 and 2.5 presented here shows strong regional patterns as the large number of catchments are divided into a small number of groupings. This reflects the fact that the co-evolution of catchments with the climate and the underlying geology is reflected in both the vegetation and the soils in which they grow. Thus the humid east coast with its steady supply of rainfall during the year has the ability to support large areas of deciduous forest, while the drier inter-continental regions with more variable rainfall tend to support more grassland. These differences in vegetation canopies, the underlying soils and topography, in addition to the presence or absence of a winter snowpack, all affect how a catchment stores and filters precipitation, and this is reflected in these groupings. This idea is further illustrated by another large model process class that does not appear in Figure 2.5 but does feature prominently in two of the largest catchment classes (Figure 2.4, a and b). In a large part of the Midwest, where human activity has greatly altered this co-evolution by changing the catchment storage and drainage properties (tile drainage) and even the vegetation type (agriculture), the model, based on natural processes, cannot satisfactorily

reproduce the RC for these catchments (the "No Model" class).

Table 2.1: Entropy reduction in dominant model processes sorted by class

Number of Splits	% of Entropy Remaining in Dominant Model Process
Level 0	**100.0%**
Level 1	**90.9%**
Level 2	**86.2%**
Level 3	**77.0%**
Level 4	**69.3%**
Level 5	**65.0%**
Level 6	**62.8%**
Level 7	**45.6%**

Bold signifies a significant decrease in entropy (p <1%)

In order to quantify in an objective way the findings from the visual inspection of patterns conducted so far, the connection between the catchment classes and the dominant model processes was also examined using the variability analysis described in Section 2.2.2; the result is summarized in Table 2.1. There is a significant reduction in entropy as the dominant model processes are sorted along the classification tree, i.e., the variance within each model grouping is less than that between the groups; thus the model processes display an underlying organization expressible quantitatively. This was seen qualitatively in Figure 2.4 where even the largest classes contained distinct model processes associated with them. Although the reduction in entropy was fairly consistent for each split of the tree, with entropy being reduced by about half overall, the largest drop occurred at the most bifurcated layer of the tree (Level 7). This is where many of the most seasonal catchments clustered, and these had fewer dominant model processes associated with them. In fact, this level contains the two catchment classes (ITC and ISCJ) where the model did not perform well in most cases, as well as three catchment classes where snow was a dominant factor. These catchments also cluster regionally, with most being located in the north-west-central portion of the US.

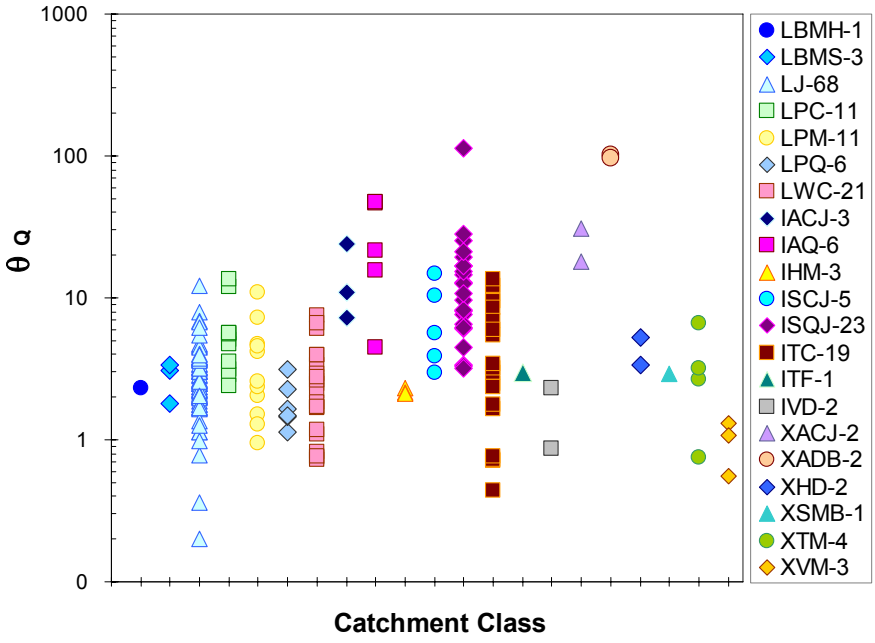

Figure 2.6: The parameter θ of the mixed gamma distribution for total streamflow grouped by catchment class and showing the number of catchments per class. (L, I, and X denote low to high seasonality in P).

2.3.2 Catchment Classes and the Parameters of the Mixed Gamma Distribution

The visual mapping of the three parameters of the mixed gamma distribution onto the catchment classes consisted of grouping the parameter values by class. For the Q duration curve, as expected, the α parameter was 0 except for classes containing arid catchments. For the scale parameter θ of the Q duration curve (Figure 2.6) catchments with low seasonality, which in this dataset are also generally more humid, have lower values (θ <20), reflecting the strong similarity in mean amount of P and Q experienced within these classes. Catchments with higher seasonality are characterized by a wider range of θ values, and within the group of intermediate to high seasonality catchments, those that are more arid (A or S in the abbreviation) tend to have a higher θ than those that are more humid. For the shape parameter κ of the Q duration curve, the pattern is not so clear. Although FDCs of arid catchments have a distinct shape, as was seen in Figure 2.5a (IAQ and XADB, both arid catchment classes), the catchment classes, which group mainly on climate properties, may not

always capture the differences in shape related to catchment properties, which are reflected in κ.

Table 2.2: Variance reduction in parameters of the mixed gamma distribution sorted by catchment class

Number of Splits	**% of Variance in Gamma Parameters Remaining**		
	α	κ	θ
Level 0	**100.0%**	100.0%	**100.0%**
Level 1	**75.0%**	96.5%	**70.1%**
Level 2	**60.8%**	93.5%	**60.5%**
Level 3	**57.8%**	87.0%	**54.8%**
Level 4	**51.9%**	79.0%	**43.2%**
Level 5	**46.4%**	75.3%	**40.2%**
Level 6	**45.8%**	74.8%	**39.8%**
Level 7	**44.9%**	72.2%	**38.8%**

Bold signifies a significant decrease in entropy (p <1%)

As before, the visual mapping of the connection between the catchment classes and the three parameters of the mixed gamma distribution was followed by a quantitative variability analysis; the result is summarized in Table 2.2. These results show that with each step down the tree the overall variability of all three parameters decreases, and thus these, like the dominant model processes, are explained by the classification tree in quantitative terms. However, not all parameters sort equally well; both α and θ show significant reduction in variance while κ does not. This may be explained by the basis for the classification scheme: the four criteria are mainly climate-oriented, as is α, and they utilize the RC, which reflects the average values of P and Q over a long time period, as does θ Where variance reduction *is* significant, it is greater in the first few splits on the tree than at the later branches. For α, this is because α is nonzero when there is no P or Q; the larger the α, the more intermittent the rainfall or streamflow. The first two branchings of the tree immediately split off a large portion of humid catchments, where a ≈ 0 for Q and Q_u, and is small for P and Q_f, from the rest, thus providing an immediate, and large, reduction in variance. From the visual mapping, these same catchments also experienced very low seasonality and had very small ranges of θ values; thus splitting them off from their more variable counterparts provided

immediate organization with respect to θ.

2.3.3 Dominant Processes and the Parameters of the Mixed Gamma Distribution

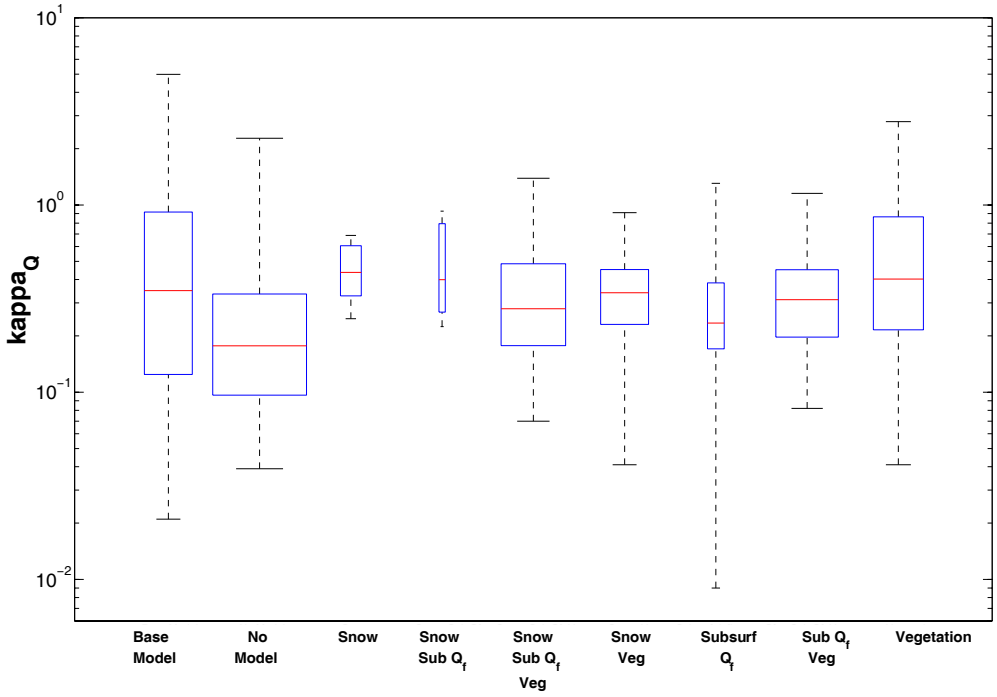

Figure 2.7: The parameter κ of the mixed gamma distribution for total streamflow grouped by dominant model process. Note that box width reflects number of catchments in each group.

The last part of the three-way mapping was the connection between the dominant model processes and the parameters of the mixed gamma distribution. For α, the story is the same as outlined in the previous section. For the shape parameter, the process class "Vegetation" showed a wider variability of κ values relative to most of the other classes, notably large classes such as "Snow, Vegetation", "Snow, SubQ_f, Vegetation", and even "SubQ_f, Vegetation" (where "SubQ_f" refers to subsurface-influenced Q_f) as seen in Figure 2.7 Vegetation as the sole dominant process seems to produce a large variability in the shape parameter κ. Although vegetation processes were lumped together into one model process class, different kinds of vegetation (grassland, forest) affect the streamflow RC in different ways – the

phenology signature of forested catchments, for example shows up clearly as a significant drop in summer streamflow. On the other hand, snowmelt affects the RC in one way – a sharp peak in spring. While not visually obvious when plotting the FDCs, these variations may be showing up in the shape parameter of the fitted FDC as a wider range of κ values. However, when snow and/or subsurface-influenced Q_f share dominance with vegetation, the variability seems to be lessened. From our inspection of the shapes of the FDCs in Figure 2.5, this seems to be borne out, at least for "Snow, SubQ_f, Vegetation" and "SubQ_f, Vegetation", and may be caused by the presence of subsurface-influence Q_f. We also saw from Ye et al. (2012) and Coopersmith et al. (2012) that snow-dominated catchments share similar climate and physical properties; in addition, the presence of snow in heavily vegetated areas further groups the catchments by climate and geology/topography, as the majority of these catchments are found in the northeastern US. This may explain in part the smaller variation in κ values for process classes containing snow and vegetation. However, process classes with snow but without vegetation contain many fewer catchments, and thus a more quantitative examination is necessary to determine if the smaller variability in the κ parameter is significant. For θ, the trend was not so clear, likely due to the wide range of climates (and therefore average P and Q) experienced within some model process classes.

Table 2.3: Variance reduction in the parameters of the mixed gamma distribution grouped by dominant model process

Grouping	**% of Variance Remaining**		
	α	κ	θ
Ungrouped	100.0%	**100.0%**	100.0%
Model Process	95.6%	**92.0**%	93.3%

Bold signifies a significant decrease in entropy (p <1%)

As shown in Figure 2.7, some process classes contained fewer catchments than others, and so a visual assessment alone is insufficient to analyze differences between classes. Thus the connection between the dominant model processes and the three parameters of the mixed gamma distribution was analyzed quantitatively using the variability analysis described in Section 2.2.2; the result is summarized in Table 2.3. As before, the results show a decrease

in variance when the parameters are grouped compared to when they are not. Interestingly, when the parameters are grouped by dominant model process, the opposite is seen compared to when they are grouped by catchment class. Here κ shows statistically significant variance reduction whileα and θ do not; thus the shape parameter κ is better-described by the dominant model processes while the scale parameter θ and α, the fraction of no rainy or flow days, are better described by the catchment classification system. This may be because the classification system is based in large part on the climate – the amount and timing of water available to a catchment – while the dominant model processes are more closely related to catchment function, since they explicitly take into account the effect of interception, phenology, and subsurface-influenced Q_f.

2.3.4 Synthesis: the Slope of the FDC

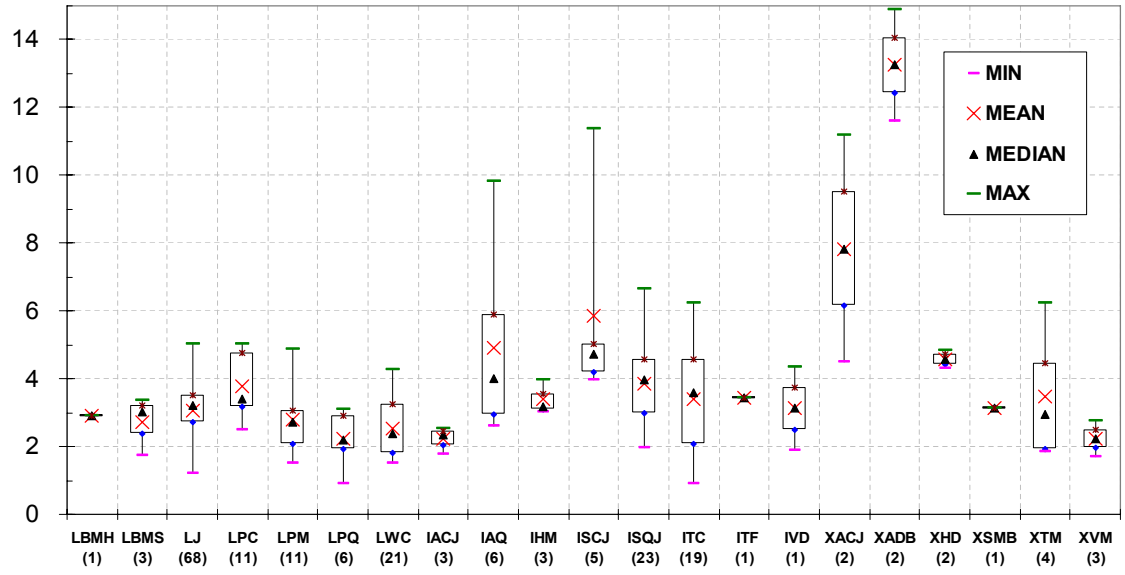

Figure 2.8: Total streamflow FDC slopes grouped by catchment class.

As was previously discussed, there is more than one way of quantifying the FDC: the fitted parameters of the mixed gamma distribution (α, κ, θ) or the slope of the FDC. However, the three parameters describe the entire FDC, while the slope as defined here only describes the

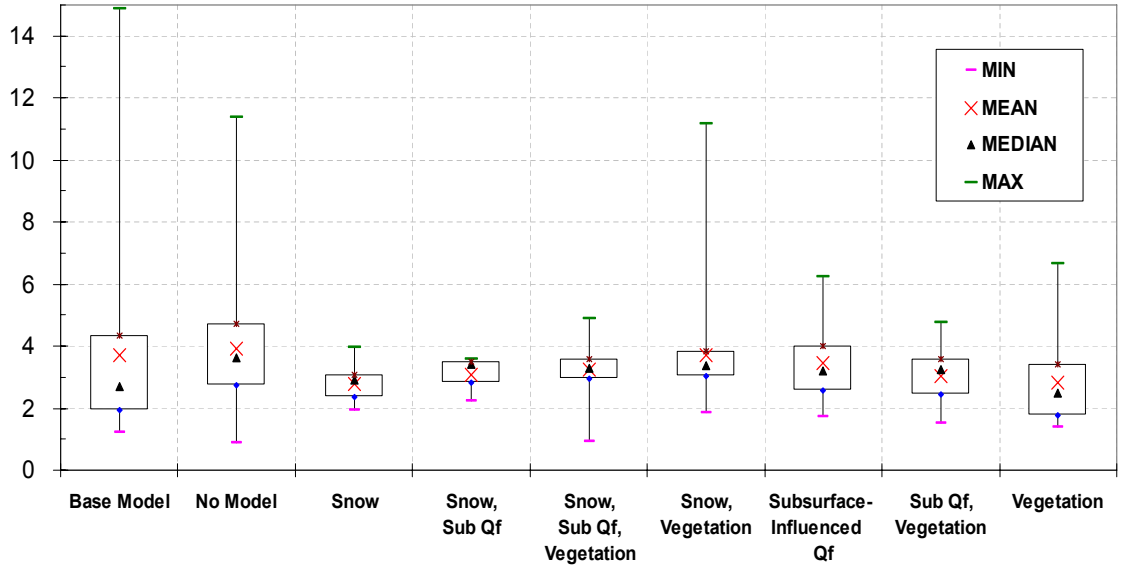

Figure 2.9: Total streamflow FDC slopes grouped by dominant model process.

middle limb – the bridge between the high flows and the low flows. By focusing on one limb of the FDC at a time, more specific controls may be identified, since the analyses in previous sections were not able to fully untangle the climate from the catchment controls. In this paper we only examine the slope the middle limb of the FDC, as it is closely related to the RC, on which the modeling and classification studies were based; the upper and lower limbs are left for future work. The slopes of the FDC for Q and Q_u were grouped first by class (Figure 2.8) and then by dominant model process (Figure 2.9). The box plots provide a visual assessment of the variability of the magnitudes of the slope of the FDC within each grouping: the size of the box represents the 25th – 75th percentile, while the whiskers show the minimum and maximum values. Because of the way the slope is calculated, catchments with ephemeral streams or those for which the streamflow becomes very small at higher percent exceedance (i.e. drier catchments) will have very large slopes. This is clearly illustrated by the class XADB (high seasonality, arid, wetter winters) in Figure 2.8 which consists entirely of arid

catchments in southern California.

In general, the FDCs of classes with low seasonality (those with L in their designation) tended to have smaller slopes and less variability in slope values than did those classes with higher seasonality (those with I or X in their designation). The FDCs of more arid catchments (those with S and A in their designation) with higher seasonality tended to have steeper slopes. In general, the majority of catchments with snow as a dominant process tended to have less variability in slopes than did those without (Figure 2.9). However, from previous examination (Figure 2.7) we see that some of these process classes are quite small in size compared to others. Nevertheless, when examining the four largest process classes aside from "Base Model" and "No Model", the majority of the catchments in "Snow, Sub Q_f, Vegetation" and "Snow, Vegetation", showed more narrow ranges of FDC slopes than did "Sub Q_f, Vegetation" and "Vegetation". These results mirror those of the visual mapping done in previous sections, where low seasonality reduced variability in θ values, and where snow processes reduced the variability in κ values.

Table 2.4: Variance reduction in FDC quantiles (a) and slopes (b) sorted by catchment class

Number of Splits	**% of Variance Remaining**			
	(a) FDC		**(b) FDC Slope**	
	Q	Q_u	Q	Q_u
Level 0	**100.0%**	100.0%	100.0%	**100.0%**
Level 1	**85.6%**	89.8%	84.2%	**81.9%**
Level 2	**68.2%**	85.1%	76.0%	**71.6%**
Level 3	**63.2%**	83.3%	74.8%	**68.2%**
Level 4	**54.2%**	78.9%	68.4%	**61.9%**
Level 5	**46.7%**	74.2%	63.3%	**57.0%**
Level 6	**46.8%**	74.1%	63.2%	**58.2%**
Level 7	**46.4%**	72.5%	62.1%	**57.2%**

Bold signifies a significant decrease in entropy (p <1%)

As with the three-way mapping, the visual assessment was followed by a variability analysis. Both the slope of the FDCs and the FDC quantiles were sorted on the catchment classification decision tree. When the FDC quantiles were sorted along the tree, there was

a monotonic decrease in variance for all duration curves, yet as before, the most variance was explained in the first two levels. The result for the Q FDC was significant (Table 2.4, a), while that for Q_u was not. When the slopes of the Q and Q_u FDCs were sorted along the tree, the same pattern of early reduction of variance leveling off with further branching was repeated, but this time, the slopes of the Q_u FDC showed a more significant reduction in variance than those of Q (Table 2.4, b). The overall pattern seen here is similar to the reduction in variance obtained by sorting the gamma parameters on the classification tree; there too the decrease in variance (where significant) was greater in the early branchings than in the outer branches. This may again be explained in part by the way the classification tree branches – the early splits remove a large portion of the low-seasonality, humid catchments from the distribution, which as visually mapped in Figure 2.4c and d, have very similar FDCS, with the majority of the differences between the FDCs in each class seen at flows occurring >66% of the time – i.e. the lower limb of the FDC.

Returning to the conceptual model presented in Section 2.2 of this paper, we have now analyzed the catchment response indicator, in the form of the FDC, from both the perspective of the entire FDC (parameters of mixed gamma distribution) and the middle limb of the FDC/ regime curve (slope). What remains is to examine more closely the relationship between the FDC and the climate and catchment properties that define its shape. The Q FDC was plotted against both the seasonality index (SI), representing the climate, and the baseflow index (BI), representing the catchment (although we should note that climate factors into BI as well). As an additional layer of analysis, the slopes of the FDCs were grouped first by catchment class (Figure 2.10) and then by dominant model process (Figure 2.11) in order to determine if any clustering becomes evident, which could help bring out which process governs which part of the FDC and aid in regionalization.

From both Figure 2.10 and Figure 2.11, it can be seen that there is a much stronger relationship between the slope of the Q FDC and BI than with SI, although this seems to break down at higher slopes (in arid catchments other processes may be affecting slow flow). The catchments with the highest slopes all belong to arid or semi-arid classes with mid- to high climate seasonality. The dominant model processes for these same catchments vary between No Model (in some arid catchments, the simple model was not complex enough

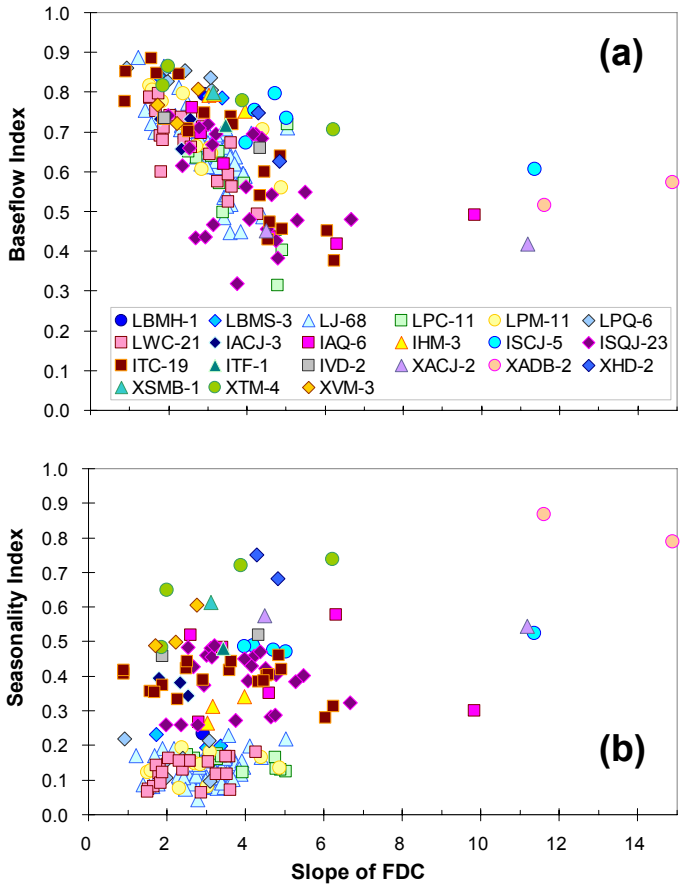

Figure 2.10: Relationship between baseflow index (a) and seasonality index (b) and total streamflow FDC slopes (grouped by catchment class).

to capture the regime curve) and Base Model (here the climate seasonality is the only complexity needed). Of these arid catchments with a large FDC slope are two that at first glance may seem at odds with the aridity of the region: one is dominated by vegetation processes, and the other by both snowmelt and vegetation processes. The first catchment is on the Caney River in Kansas, one of the Great Plains states, and closer inspection reveals a lush riparian corridor. Due to the aridity of the region, the stream may become dry in the summer months, thus causing the large FDC slope. In dry regions, woody riparian vegetation are often phreatophytes, which, by tapping into shallow groundwater, draw down the water table and in the case of a gaining stream can reduce streamflow, thus contributing in part to the steep slopes seen in these FDCs. The second catchment is on the Bad River in

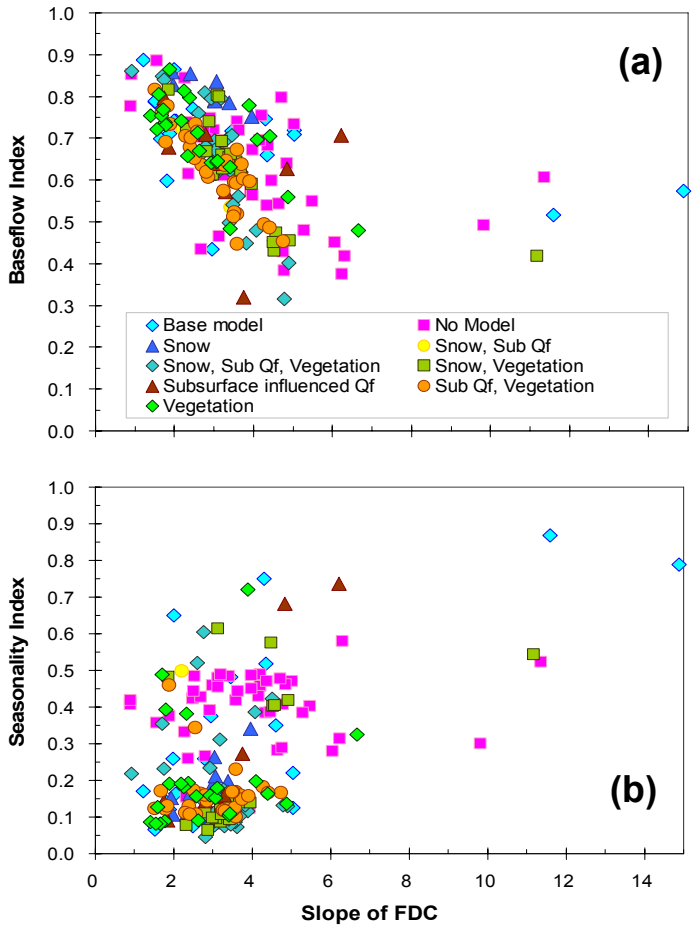

Figure 2.11: Relationship between baseflow index (a) and seasonality index (b) and total streamflow FDC slopes (grouped by dominant model process).

South Dakota, and is a snow-fed tributary of the Missouri River that flows in part through a national prairie grassland. The plot of the slope of the Q_u FDC (not shown here for brevity) follows that for the Q FDC, except shifted slightly to the left, as the slope of the Q_u FDC is generally less than that of Q. This is likely due to the extreme high flows present in the Q FDC pulling the curve upwards, thus steepening the overall slope.

2.4 Discussion

2.4.1 Comparative Analysis: Regionalization

In all three studies – statistical, modeling, and classification – one common theme emerged: patterns, both spatial and empirical, pointed toward an underlying structure or control. The current work has attempted to draw connections between the results of previous work to elucidate the nature of these controls. This pooling of information is the essence of synthesis; examining the information gained from several analyses leads to a deeper understanding of the questions, which can then lead to better answers, as well as guide the direction of future work. Our four-part study has been comparative in nature, encompassing nearly 200 catchments with a wide gradient of climate and catchment properties. This process of learning from different places (a Darwinian approach) runs counter to the current general ethos of hydrology, which is to learn from an individual place in great detail (the Newtonian approach). Our approach has been to step back and look at this issue holistically through a combined approach, where information gained from detailed studies of individual catchments augment the information gained from the study of a population of catchments. By examining the entire dataset across a population of catchments, we have identified regions of homogeneity, and have attempted to classify them in different ways, as well as make connections between them in order to clarify how catchments behave in different locations, thus partitioning the nation into regions of similarity. By determining the underlying processes within a region, and in turn how they might vary by region, we gain understanding of the physical controls and in this way step closer to the ultimate goal, the prediction of the FDC, in any ungauged basin.

The regional patterns highlighted in these four papers led to the groupings that have been explored both qualitatively and quantitatively in this synthesis analysis. We have shown that the variance within these groups, both in the three-way mapping and in the FDC slope analysis, was less than that between groups, thus objectively demonstrating that these groupings are meaningful. This is consistent with the idea that classification and catchment grouping is useful for estimation of the FDC in ungauged catchments, particularly in large

areas (see Sauquet and Catalogne (2011)). Of note is one region that was highlighted in all of our studies, particularly the modeling and catchment classification: the Midwestern US. This is a mid-continent region of transition to higher climate seasonality and aridity, and is also this country's main agricultural region. While the MOPEX catchments are ostensibly minimally impacted by human behavior (Duan et al., 2006), recent work by Wang and Hejazi (2011) has shown that this is not always the case. These human alterations have, in part, become dominant processes themselves, as illustrated by the simple model's inability to perform satisfactorily in this part of the US, and thus must be taken into account in any regionalization studies with significant anthropogenic activities.

The synthesis analysis of these groups showed that while the classification tree described well the scale parameter θ, and the zero exceedance parameter α, the dominant model processes better described the shape parameter κ. Since the model processes focus more on catchment function, this gives first-order physical meaning to the shape parameter. However, there was no unique value of κ for each model process, which points to the need for a more complex model than we use here, such as the process-based model of Carrillo et al. (2011) and others like it, that can capture catchment properties in greater detail and thus bring out their controls on the FDC. Since the climate and catchment function are so closely linked through geologic time by co-evolution, more specific properties of the catchment, such as topography, soil types, etc, all of which determine in part the filtering of infiltrated precipitation and thus the FDC at the outlet, may need to be included. Thus, with the catchment more explicitly defined in the model, the effects of the catchment itself may be separated from the effects of the climate it experiences. In addition, the parameter fitting itself was strongly affected by the tails of the FDC. Parameter values containing more close fits to the tails would likely produce stronger synthesis relationships with the model processes.

Further analysis into the slope of the Q FDC revealed that the baseflow index was a stronger control on it than was the seasonality index. The slope of the FDC as defined here is related to the RC, which is a function of long-term catchment response, and this would confirm the idea that the middle limb of the FDC is related to the RC. Clustering of classes and model processes was seen in the scatter plots (Figures 2.10 and 2.11); catchments in the region where snowmelt is the dominant process tended to all have higher BI, while the

vegetation process dominated regions explored a larger range of BI. Combining the results of this study and the empirical one, we see that catchments with larger κ values tended to also have larger BI (a greater proportion of Q exits as Q_u) and flatter Q and Q_u FDC slopes. These results seem to indicate that catchment properties rather than the more time-dependent climate properties are the dominant controls on the FDC. Climate seasonality does affect the catchment response as the classification study highlighted, and to some extent the shape of the FDC, as the empirical study highlighted, but the control is secondary to the catchment properties, as evidenced by the stronger relationship of the slope of the FDC to BI compared to SI, and the more significant variance reduction of the Q_u FDC compared to the Q FDC when sorted on the classification tree. Seasonality implies timing; the FDC removes the time element, and by doing so may also provide a first order separation of climate influence from that of catchment properties.

That the relationship is stronger with BI may be due to the way the slope of the FDC is calculated. A slope obtained using Eq. (2.2) represents the middle limb of the FDC, which itself represents the long-term, averaged response from the catchment. Since the baseflow index also is a measure of these processes, and as such is governed more by the topography and other earth processes than by climate processes, this result confirms the hypothesis of Yokoo and Sivapalan (2011) that the middle limb of the FDC is a function of the RC. This relationship of the shape of the FDC with groundwater-related measures has also been seen previously in the literature, in studies not limited to the US. For example, Claps and Fiorentino (1997) use BI to regionalize the FDC in Basilicata, in southern Italy. Similarly, in the UK the HOST classification of soils is used as a surrogate for the geology (Holmes et al., 2002). Sauquet and Catalogne (2011) found that the catchment yield and the percentage of impermeable substrata (which are more or less related to the geology) are among the most important explanatory variables to regionalize the FDC in France.

Regarding seasonality, which showed a weaker relationship to the slope of the FDC, not many papers use it explicitly to regionalize the FDC. One example is Sauquet and Catalogne (2011) who use the seasonality index to group French catchments into regions of different regression applications. On the other hand, the seasonality is implicitly accounted for in other works. For example, in Arora et al. (2005) and Ganora et al. (2009) the mean catchment

elevation is used to discriminate between regions with different FDCs in Nepal and Italy respectively, and it is argued that elevation is a proxy for distinguishing areas more or less influenced by snow mechanisms (and therefore different seasonality). Li et al. (2010b) found that the Leaf Area Index is correlated to the shape of the FDC in south-eastern Australia; vegetation impacts ET differentially between summer and winter, increasing the slope of the FDC and the probability of having zero flow for some portion of the year, similar to what was seen in some of the more arid regions in the US. These implicit seasonality effects were also seen in the MOPEX catchments when vegetation-related dominant model processes led to higher variability and steeper slopes in the FDC. Thus the findings of this study, based solely on catchments in the continental US, correlate with findings from a more global perspective and further highlight this co-evolutionary relationship between climate and catchment properties.

2.4.2 Comparative Analysis: Regime Curve and the Flow Duration Curve

The aim of this study was to determine the underlying physical controls on the FDC. However, another, less expected, theme that kept surfacing throughout the entire study, was that the differences between catchment responses are more readily seen in the RC than in the FDC. These differences between catchments appear more obviously, and the reasons for those differences are far more straightforward, in the RC in part because the temporal pattern remains intact, while in the FDC this information is removed. From both the catchment classification study (Coopersmith et al., 2012) and the process modeling study (Ye et al., 2012), it was shown that many of the processes controlling the catchment response are time-dependent, for example, the seasonality of precipitation or the vegetation processes of tree leaf-out and leaf-drop. In addition, it is not uncommon for more than one hydrologic control process to be at work in a given catchment; thus one complicating factor in relating the regional patterns seen in these studies to the physical controls on the FDC is that the patterns obtained for both the dominant model processes and the classification system were based on the RC, not the FDC. The guiding principle behind these two studies was that the RC and middle limb of the FDC are related to each other, and we have seen evidence

of this in both this paper and in the statistical study (Cheng et al., 2012). To further understand this relationship between the RC and the FDC, we have again approached it both qualitatively, by examining RCs and FDCs from a few selected catchments in the study, and quantitatively, by comparing the slopes of both the FDC and the RC plotted as an FDC.

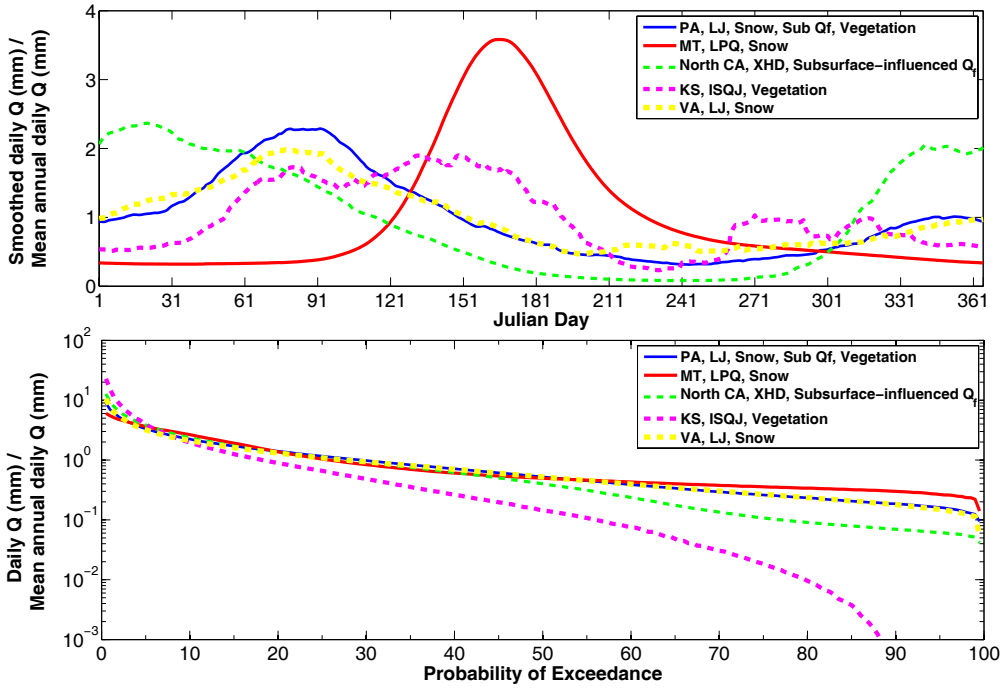

Figure 2.12: Smoothed regime curve of total streamflow (*top*) and corresponding flow duration curve (*bottom*), both at the daily scale and normalized by mean annual daily flow. (Legend description is given as US state abbrev., catchment class, dominant model process).

Qualitatively, the information about process controls on the catchment response contained within the FDC is more subtle, since removal of the time element filters out what this work has shown to be crucial information. This might lead one to think that two catchments with different regime curves could have the same FDC. However, as Figure 2.12 shows, this is not always true; there is still some information to be gained from the FDC. For the catchment in Montana (MT), when looking at the RC, the peak in streamflow caused by snowmelt is quite prominent, but for the rest of the year, Q is relatively constant. This can be seen in the FDC of MT as a fairly flat slope through most of the curve, with a slight uptick at the rarely-exceeded (high flows) end, as in this catchment the highest flows generally

correspond to the annual snowmelt events. Contrast this with the semi-arid catchment in Kansas (KS), where the streamflow varies a great deal throughout the year. This is seen in the FDC of KS as a much steeper slope overall, with the aridity of the catchment indicated by the curve tending to zero flow before 100% exceedance. Thus catchments with very different RCs also have different FDCs, although this difference is much less pronounced. Catchments with very similar RCs have even more similar FDCs, as illustrated by the two forested catchments in Pennsylvania (PA) and Virginia (VA); here the seasonal pattern of leaf-out in the spring and leaf-drop in the fall are seen in the RC as a decrease in streamflow from about May to October, even though the precipitation for the east coast of the US is fairly constant all year. This phenology signature is less clear in the FDC, where the two catchments are nearly indistinguishable from one another – except at the extreme ends – and not substantially different from the MT catchment described earlier. In spite of this, some differences in the shape of the FDCs can be related to possible physical controls; the arid catchment in Kansas being a good example of this. Note that the major differences in these example FDCs are in the tails; this is where the story of the FDC may lie, and it is possible that an analysis of extremes may illuminate this further. Also, this example only highlighted the Q FDC; it may also be that the FDCs of the components Q_f and Q_u, as well as their tails, may further elucidate the physical controls on the FDC.

The qualitative visual analysis of Figure 2.12 highlighted the information contained in the RC that does not readily appear in the FDC. However, by examining both signatures in a quantitative fashion, it becomes clear that there is information in the FDC that is not contained in the RC. When the smoothed daily RC is normalized by its mean, then sorted and plotted as an FDC, the slope of the resulting curve can be calculated using Eq. (2.2), but with the highest and lowest values instead of the middle third. Figure 2.13 shows these slopes plotted against the slopes of the middle third of the FDC, and grouped by catchment class. In general, the slopes of the FDCs are larger than those of the RC FDCs. This is because the FDC contains information about the extreme flows (both high and low) that a catchment experiences, and these pull the middle limb of the FDC in opposite directions, thus increasing the slope. The daily RC, being a long-term average of all these flows, has the extremes smoothed out; thus the slope of an RC FDC is much

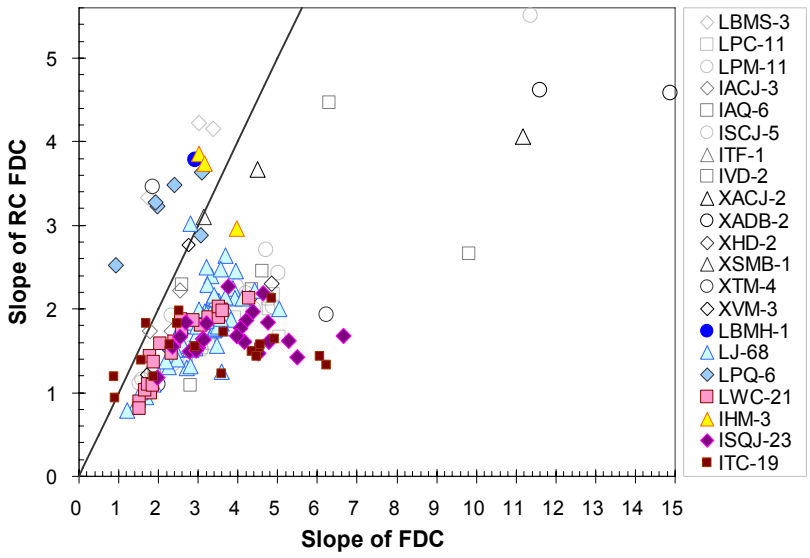

Figure 2.13: Relationship between the slope of the FDC and the slope of the RC when plotted as an FDC (grouped by catchment class). A 1:1 line is provided for reference.

lower than the corresponding FDC middle limb. By grouping these slopes by catchment classes, we see some interesting regional patterns emerge. The catchments where the RC FDC slope is the same or larger than the corresponding FDC cluster mainly in the low-to-mid seasonality, more humid catchments in the north west central US (e.g. LBMH, LPQ, IHM). This region is dominated by snowmelt processes, as illustrated by the MT catchment shown in Figure 2.12, and this is seen quantitatively in both the RC and the FDC through their slopes. Catchments in the low-seasonality, humid east coast (e.g. classes LJ and LWC) tended to have RC FDC slopes closer to that of their corresponding FDC than did the intermediate-seasonality catchments in the mid-continent region (e.g. ISQJ and ITC). In the eastern, deciduous-forested catchments, the spring flow peak followed by the sharp decline in streamflow during the summer months, as illustrated by the PA catchment in Figure 2.12, serves to mimic the extremes seen in the FDC tails, thus steepening the RC FDC somewhat. This, combined with the flatter overall slope of FDCs in humid regions, brings the RC FDC slope closer to the slope of the corresponding FDC. However, this clear Spring peak/Summer low flow regime behavior is generally not seen in the RCs of catchments in the Midwest, which are dominated by grassland and more seasonal precipitation. Because the

averaging processes inherent in producing a regime curve further smooth out the differences between the extreme flows, RC FDC slopes from catchments in this region deviate even further from their corresponding FDC's slope.

2.5 Conclusions

Each aspect of this four-pronged analysis revealed regional patterns that led to insights into the first order controls on the flow regimes and flow duration curves in nearly 200 catchments across the Continental US and laid the foundation for a framework that can be built upon in the future. The empirical study of the FDC (Cheng et al., 2012) showed that geology and landscape properties of the catchment have a stronger influence on the shape of the FDC than do climate properties, but the scatter in this relationship shows that even here climate is exerting some influence. Regional patterns in the values of the parameters of the mixed gamma function led us to classification as a way to organize and explain the underlying controls. Guided by the hypothesis that the RC provides the backbone to the shape of the FDC (Yokoo and Sivapalan, 2011) we classified these same catchments in terms of their regime behavior. This classification was done both empirically (Coopersmith et al., 2012) and with the aid of process-based modeling (Ye et al., 2012), and both methods showed that climate (seasonality, followed by aridity) were the dominant controls on the regime behavior across the continent, overriding the catchment controls of topography, soils, etc. Quantitative analysis performed in this synthesis paper showed that both the process and empirical classes, which are based on the RC, were connected to the broad properties of the flow duration curve (the slope of the middle limb and the parameters of the mixed gamma function). However the regime behavior captures different aspects of variability compared to what is reflected in the FDC. While the regime curve smooths out the high and low flow behaviors, which together can impact the slope of the middle limb of the duration curve, the FDC retains this information in the upper and lower tails. Thus in the FDC, time information (seasonality) is lost, but information about the extreme flows is gained.

The visual mapping done here showed that within each regime class, both process and empirical, there was a great deal of heterogeneity in the shapes of the FDC which was most

noticeable in the tails. This emphasized a finding of the first study, where the statistical fitting was strongly affected by the tails of the FDC. While the upper tail is most affected by the precipitation intensity, the lower tail is affected more by catchment properties of landscape, soils, geology, etc than by climate. Thus further analysis of the tails of the FDC, especially the low flows, may separate the physical controls on the shape of the FDC more effectively. In moving forward, the modeling study must also be enhanced to incorporate these high and low flow processes. The similar regions delineated in this study, based on the regime curve, can be used to "zoom in" on a particular class, guiding the application or development of more detailed models to bring out the differences in the underlying physical (as well as anthropogenic) controls at work in each region. Under this framework, the coupled soil-water-plant growth analytical model of Feng et al. (2012) or the detailed process-based model of Carrillo et al. (2011) are the kind of models that would be better suited to fully separate the catchment controls of the FDC from the climate controls. We leave the pursuit of this exploration to future research.

Acknowledgments

The work presented in this paper was carried out as part of the NSF-funded project "Water Cycle Dynamics in a Changing Environment: Advancing Hydrologic Science through Synthesis" (NSF grant EAR-0636043, M. Sivapalan, PI), also the NSF project "Understanding the Hydrologic Implications of Landscape Structure and ClimateToward a Unifying Framework of Watershed Similarity" (NSF Grant EAR-0635998, T. Wagener, PI). Special thanks are owed to Matej Durcik of SAHRA (University of Arizona) for providing the version of the MOPEX dataset used in this study.

Chapter 3

Comparative analysis of hydrologic signatures in two agricultural watersheds in east-central Illinois: Legacies of the past to inform the future

Abstract

Historically, the central Midwestern US has undergone drastic anthropogenic land use change, having been transformed, in part through government policy, from a natural grassland system to an artificially-drained agricultural system devoted to row cropping corn and soybeans. Current federal policies are again influencing land use in this region with increased corn acreage and new biomass crops proposed as part of an energy initiative emphasizing biofuels. To better address these present and future challenges it is helpful to understand whether and how the legacies of past changes have shaped the current response of the system. To this end, a comparative analysis of the hydrologic signatures in both spatial and time series data from two central Illinois watersheds was undertaken. The past history of these catchments is reflected in their current hydrologic responses, which are highly heterogeneous due to differences in geologic history, artificial drainage patterns, and reservoir operation, and manifest temporally, from annual to daily time scales, and spatially, both within and between the watersheds. These differences are also apparent from analysis of the summer low flows, where the more tile-drained watershed shows greater variability overall than does the more naturally drained one. In addition, precipitation in this region is also spatially heterogeneous even at small scales, and this, interacting with and filtering through the historical modifications to the system, increases the complexity of the problem of predicting

This work has been published as: Yaeger, M. A., Sivapalan, M., McIsaac, G. F., and Cai, X.: Comparative analysis of hydrologic signatures in two agricultural watersheds in east-central Illinois: legacies of the past to inform the future, Hydrol. Earth Syst. Sci., 17, 4607-4623, doi:10.5194/hess-17-4607-2013, 2013. All figures, tables and data were created by Mary Yaeger unless otherwise indicated.

the catchment response to future changes.

3.1 Introduction

The Renewable Fuel Standard (RFS), a provision of the US Energy Policy Act of 2005, mandated 7.5 billion gallons of renewable fuels by 2012, and was subsequently expanded to require 36 billion gallons by 2022. This mandate has created and increased the demand for biofuels, leading to an increasing demand for biofuel refinery feed-stocks. Since current biofuel production consists mainly of corn-based ethanol, this has led to a rapid expansion in corn planted area, the majority of which is found in the Corn Belt, a fertile, humid region in the Midwestern US, comprising portions of the Upper Mississippi River Basin as well as portions of the Ohio River Basin. Recent studies (Donner et al., 2004; David et al., 2010), however, have pinpointed this region as the source of the majority of nitrate exported to the Gulf of Mexico, which is a cause of the large hypoxic zone at the mouth of the Mississippi River. Amid concern over increasing nitrate levels resulting from increased corn crop fertilization (Donner and Kucharik, 2008), the RFS has recently been expanded to now include a mandate for second generation biofuels, providing further motivation for research into alternative feed-stocks for refining biofuels. Of these, perennial biomass crops such as *Miscanthus giganteus* and switchgrass, grown for cellulose-based ethanol production, have shown much promise. Both field experiments (McIsaac et al., 2010; Smith et al., 2013) and watershed-scale modeling experiments (Ng et al., 2010) have shown that less nitrate is exported from *Miscanthus* compared to conventional corn crops. However, recent field experiments (McIsaac et al., 2010; Hickman et al., 2010) and canopy-scale modeling studies (Le et al., 2011) have also shown that water usage by *Miscanthus* is significantly greater than that of corn or soybeans. If large-scale planting of these biomass crops is to be sustainable, it is essential to understand how watersheds as a whole might respond to this change and be aware of possible negative outcomes.

This switch to perennial grasses on the scale required to meet the cellulosic ethanol demand set forth in the RFS may cause a major change to the agriculture which then will impact the hydrologic processes in the region. What the future hydrologic responses to biofuel

development in the region will be remains a concern. To address this question, a regular approach is to develop a hydrologic model of the region, calibrate it to observed data, and then run scenarios based on the proposed changes. Given that this region has experienced similar major changes in the past, we suggest that before commencing such modeling efforts, much can be learned from a detailed analysis of historical changes to this region that led to the current conditions. Once the history of a watershed is known, more recent data can be analyzed to find the hydrologic signatures of these past changes and this knowledge can then be used to inform the modeling process. The purpose of this study, then, is to examine the legacy of past changes in two typical Corn Belt watersheds to gain a deeper understanding of the watershed response in order to better predict the response under the proposed changes.

This paper is comprised of four main sections. In the first, we have introduced the motivation for the study; the second provides a historical perspective of the natural and anthropogenic changes already imposed on this region as well as the environmental and anthropogenic responses to those changes. Section 3.3 describes the data used, spatial characteristics of the study watersheds, and the methods used in this study; the results of this comparative analysis are presented in Section 3.4. In Section 3.5 we discuss the findings and how they can be applied to the problem of predicting future changes. We conclude with a discussion of the limitations of using observed data for the purposes of prediction and a preview of future work that can overcome these limitations.

3.2 Historical Impacts and Environmental Feedbacks

The central Midwestern US presents rich examples of both natural and anthropogenic changes and corresponding environmental feedbacks. Each of these changes left its imprint on the geology, hydrology, and vegetation of this region, and these signatures can be found throughout the historical record and in the data. If we consider the region, at least from human settlement onward, as a coupled human-nature system, its history can be visualized as a series of impact/feedback loops that spiral forward in time. The proposed crop changes, then, would be the most recent in a region which has undergone many such changes in the past, often to the detriment of the surrounding environment. Through analysis of past

human impacts and environmental responses we can gain a better understanding of the dynamics of this coupled system, and therefore make more informed predictions about the responses to impacts of the new biofuels crops.

3.2.1 Pre-European Settlement to 1850: Wet Prairies

Before humans became established in the Midwest, the region was subject to natural impacts that influenced the co-evolution of soils and vegetation and resulting hydrologic conditions. In the last 1.6 million years, much of the region, including most of the state of Illinois (IL) was at some point covered by glaciers, and in some places, more than once. As the glaciers retreated, land cover became more forested as tree species from the south and east migrated to the region through seed dispersal, although further climate fluctuations, aided by fire, provided natural disturbances that kept the system in flux between grassland and forest (Whitney, 1996). Warmer, drier climate periods favored grassland development, while forests tended to expand in cooler, more humid climate periods where topography allowed, since underlying geology also played an important role in the vegetation and soils that co-evolved in the Midwest. The most recently glaciated regions were dominated by flat or slightly depressed areas overlaying clayey deposits that collected spring precipitation and remained saturated well into the summer, thus preventing establishment of woody vegetation such as trees, which are less tolerant to prolonged waterlogging. The natural condition of this region, therefore, was a shifting mosaic of grassland and forest that changed in response to climate fluctuations and fire frequency under the constraints of soil properties (Whitney, 1996; Prince, 2008; Woodhouse and Overpeck, 1998).

There is strong evidence that the first major human impacts to this region were not due to the European settlers but rather to the extensive use of fire by early humans to modify the landscape by suppressing forest in favor of grassland (Whitney, 1996; Prince, 2008). Thus when settlers began arriving in IL in the early to mid 19th century they found a region dominated by wet grassland in the north and east, with forest dominating in the south and along stream channels (Figure 3.1a). From the extensive writings of surveyors and natural historians who arrived before and with the settlers, we have some idea of the state of the

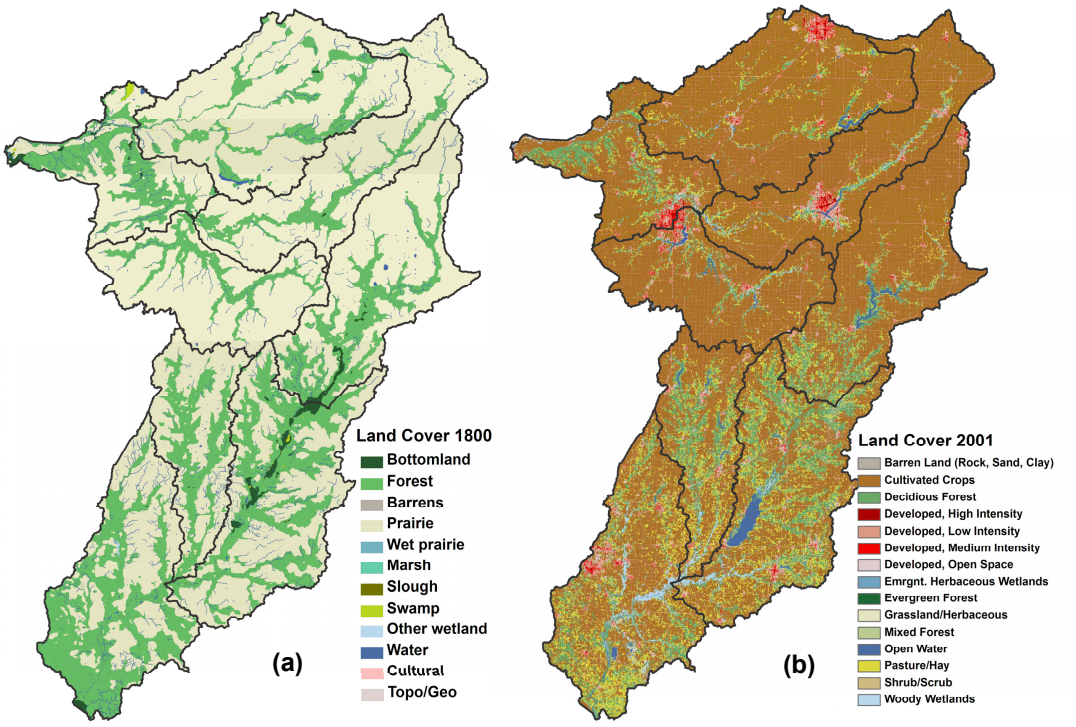

Figure 3.1: Land cover change in two watersheds in central IL: pre-European settlement land cover (a) and present-day (b).

system before the next phase of human modifications took place. Because of the geologic history of glaciation, large portions of the region, especially the uplands, tended to be flat and covered with silty clayey soils overlaying a hardpan subsoil layer; thus much of the region lacked a well-defined drainage network and the soil column was saturated in places for long periods during the year. The dominant vegetation consisted of tall grasses whose dense root systems held the fine soil particles in place, minimizing erosion from the often intense precipitation, and aiding in the retention of soil moisture in dry years. This formation of high density root mats, or sod, also posed one of the first major challenges to agricultural development by early European settlers, since plows of the time were unable to cut through it (Prince, 2008). Furthermore, precipitation falling on the vast, poorly-drained wet prairie regions would saturate the soil and then pool on the surface, from where it either infiltrated into the soil or evaporated (Prince, 2008; Jackson, 2002). In addition, wet soils provided

the necessary anoxic conditions for denitrification, which removes available nitrate-nitrogen from the shallow groundwater (McIsaac and Hu, 2004). Thus, at the time of settlement, the natural condition of most of the region was a retentive, absorbent ecosystem characterized by long residence times, and where inputs, both of precipitation and nutrients (in the form of decaying organic matter), were highly filtered and processed by the catchment (Karlen et al., 2010).

3.2.2 Impacts and Feedbacks: 1850 – 1945

From the point of view of the early European settlers, these wetland regions were useless for agriculture, and worse, were sources for malaria and other diseases (Bogue, 1951; Prince, 2008). The US Congress passed the Swampland Act of 1850, mandating that these "useless" wetlands be given to the individual states to be sold, drained, and converted to productive agricultural land. However, it was not until nearly the end of the 19th century before this was extensively accomplished. Initial efforts to drain the land by digging ditches had some success, but it was the combination of surface and subsurface drainage that proved most effective in the wet prairie regions. The passage of local laws forming drainage districts that shared the cost of improvements among all landowners, coupled with the local production of drainage tiles (Bogue, 1951) provided the impetus for the second major human impact on this region: extensive artificial drainage of mid-western wetlands and their conversion to cropland. Tile drainage lowered the water table by quickly moving water out of the soil column once it reached the level of the drains (Goswami et al., 2008); the excess water was then removed from agricultural fields by an extensive network of drainage ditches dug to connect the tile outlets to existing channels, which were often then dredged and straightened to accommodate the increased flow (White et al., 2003; Royer et al., 2006; van der Velde et al., 2010). In hydrologic terms, this greatly increased the drainage density (defined as the ratio of the total length of all stream channels to the total area drained) of the system. The end result of these changes was a switch from a retentive, poorly-drained system to a leaky, well-drained one.

With the rich soil now dry enough for crops and aided by an earlier development of

special "sod-busting" plows to cut through the prairie grass root mats (Bogue, 1994), by 1900 agriculture became the dominant land use in the region, and until about 1945, consisted of annual crops such as corn in rotation with sod-based crops such as oats or hay (Jackson, 2002). This replacement of the tall-grass prairie with agricultural crops resulted in the destruction of the grass root networks that had held the fine soils together. Freshly plowed and seeded land was highly vulnerable to rainfall erosion, especially in the humid Midwest where precipitation peaks in the late spring, and soil moisture is often near field capacity. Although soil losses were noticeable, and there were scattered attempts at conservation, many people at the time viewed the land as a vast, inexhaustible resource, and simply moved on when crop yields in one location declined (Trimble, 1985; Whitney, 1996). In this way, agricultural expansion continued westward, into a windy, semiarid climate for which the current farming practices were poorly-suited (Baumhardt, 2003). Beginning in 1931, over two-thirds of the country, including the Midwest, experienced a severe drought (Woodhouse and Overpeck, 1998); this, coupled with the effects of years of poor land management, resulted in the loss of much of the topsoil in the western Great Plains in what is now known as the Dust Bowl (Schubert et al., 2004). Although the massive environmental damage, itself one of the first major environmental feedbacks on the impacts of human agricultural activities (Cook et al., 2009; Karlen et al., 2010), was limited to the Great Plains west of the Mississippi River, it was a national economic disaster affecting the entire farm economy. It was the magnitude of the effects of the Dust Bowl that finally prompted a concerted human system response to the problem of soil erosion. In 1933, the US government formed the Soil Erosion Service (later, Soil Conservation Service) to determine the causes and extent of soil erosion in the US, and then go out and teach farmers across the country, including those in the Midwest, better soil management practices. As a result, soil erosion was significantly reduced (Turner and Rabalais, 2003; Karlen et al., 2010), with soil conservation efforts continuing to the present day.

3.2.3 Impacts and Feedbacks: 1945 – Present

The next series of major land use changes began around 1945 and have continued to the present day. In 1944 and 1945, the US government enacted two flood control acts that spurred another tile drainage development period by engaging federal agencies in land drainage work as part of their flood prevention and soil conservation missions (Wooten and Jones, 1955; Karlen et al., 2010). In addition, as tractors replaced horses, soybeans began to replace oats and hay in rotation with corn (Jackson, 2002; McIsaac and Hu, 2004), leading to the current land use configuration shown in Figure 3.1b. As large-scale row crop agriculture expanded in this region, inputs of nutrients – mainly nitrate-nitrogen and phosphorus – to the system greatly increased. The change from annual crops in rotation with sod crops to a rotation of only annual row crops decreased annual evapotranspiration (ET) by shortening the growing season, and this, coupled with increasing tile drainage density, has likely increased baseflow to local streams (Zhang and Schilling, 2006). Over this same period, annual precipitation in the Midwest has been increasing (Raymond et al., 2012), contributing to higher streamflow overall. The effect of tile drainage was to decrease soil residence times by moving water more quickly through the subsurface to the stream, thus bypassing natural biogeochemical processes that reduce the nitrate concentration in soil pore water (McIsaac and Hu, 2004; Panno et al., 2008). Furthermore, the majority of nutrient export occurs during periods of high discharge (Royer et al., 2006) when in-stream removal by de-nitrification processes is least efficient (Royer et al., 2004). In addition, wet years can result in nitrate flushing from the system, where nitrate stored in the soil during drier years is mobilized in addition to that from the current year's fertilizer application (Gentry et al., 2009). The combination of climate and land use changes and the hydrologic modification of the landscape have resulted in a marked increase of nutrient export from the Midwest region to the Mississippi River (David et al., 2001). A notable exception to this general regional pattern are those streams with large, in-line reservoirs; since reservoirs act to increase residence times, significant de-nitrification can take place, in some cases removing nearly 50% of the nitrate (David et al., 2006).

Similar to the problem of increasing soil erosion earlier in the century, the effects of in-

creased nutrient export from Midwestern agricultural watersheds initially manifested locally. Levels of nitrate in excess of US drinking water standards were frequently found in surface water bodies in the region (Smith et al., 1990; Kalita et al., 2006), including those used for municipal water supply (Keefer et al., 2010). As before, the human system response to the problem tended to be localized, although some national attention was paid. Much like the Dust Bowl, it was not until large scale environmental damage with national economic implications became evident that a concerted effort was made to respond. The appearance of a persistent, recurring hypoxic zone threatening important US fisheries in the Gulf of Mexico was a second, and more direct, major environmental feedback on the impacts of human agricultural activities (Goolsby et al., 1999; David et al., 2010). While soil conservation practices implemented after the Dust Bowl may have decreased the export of Midwestern soils in rivers and streams, they have had little effect on nitrogen export. Thus, as was the case in the 1930s, solutions for shrinking or preventing the formation of the hypoxic "Dead Zone" must involve large-scale modification of agricultural practices, this time to drastically reduce the export of nutrients from tile-drained watersheds in the Corn Belt region (Donner and Scavia, 2007; Scavia and Donnelly, 2007).

3.2.4 Present and Future Human Impacts

As the research into the causes and possible solutions to the "Dead Zone" progresses so too does the expansion of corn production in the Midwest in response to increased demands for ethanol, further complicating the problem (Donner and Kucharik, 2008; Martin, 2011). In an effort to address these feedbacks and improve water quality while at the same time meeting the new fuel demands, large-scale planting of high-yielding perennial biomass crops is being investigated. Since these crops have lower fertilizer requirements than the corn currently grown in this region for biofuels, this would effectively reduce the amount of nitrate exported to the Gulf of Mexico. However, this could be potentially detrimental to both human and environmental streamflow users because some of these plants require more water than do current crops. The lowest natural flows in this shallow groundwater-dependent region occur soon after peak of the growing season; thus, in an effort to solve the water quality problem a

water quantity problem may be created. Therefore, for large-scale biomass crop production to be sustainable, these trade-offs between water quality and water quantity must be fully understood. To this end, two watersheds, representative of typical Corn Belt catchments, were chosen for detailed analysis of their hydrologic response to human and natural impacts. We mainly focus our analysis on a few hydrologic signatures: the regime curve, flow duration curve, and the characteristics of a low flow analysis centered on the summer low flow season. The aim of this paper, then, is to take what can be learned from the long history of change and response to change in these watersheds, as evidenced in the streamflow of 1990-2011 and apply it toward predictions of future behavior. We expect that a comparative analysis of hydrologic signatures will provide insights into why these watersheds respond as they do and thus enable better predictions of how these watersheds may respond to unknown changes in the future.

Table 3.1: Description of and sources for all data used in this study

Data Description	**Data Source**	
Land cover map, 1800s, IL	IL Natural History Survey, Prairie Research Institute (contact for data file)	http://wwx.inhs.illinois.edu/resources/gis/glo/
30-m land cover, 2001; 30-m digital elevation map (DEM)	USGS[1] National Map	http://nationalmap.gov/viewer.html
STATSGO soils map, by state	USDA NRCS[2]	http://soildatamart.nrcs.usda.gov/USDGSM.aspx
Estimates of percent tile drainage, by county	World Resources Institute (WRI)	http://pdf.wri.org/assessing_farm_drainage.pdf
Daily streamflow	USGS NWIS	http://maps.waterdata.usgs.gov/mapper/index.html
Annual and monthly precipitation	PRISM Climate Group	http://prism.nacse.org/

[1]US Geological Survey, [2]Natural Resources Conservation Service,

3.3 Methods

3.3.1 Data description

Data used in the following analyses, including those used to create the land cover maps previously presented in Figure 3.1 are summarized in Table 3.1. All spatial data presented in this paper are available as part of much larger datasets that were then clipped to the boundaries of the study area in order to better compare in detail the characteristics of the two selected watersheds.

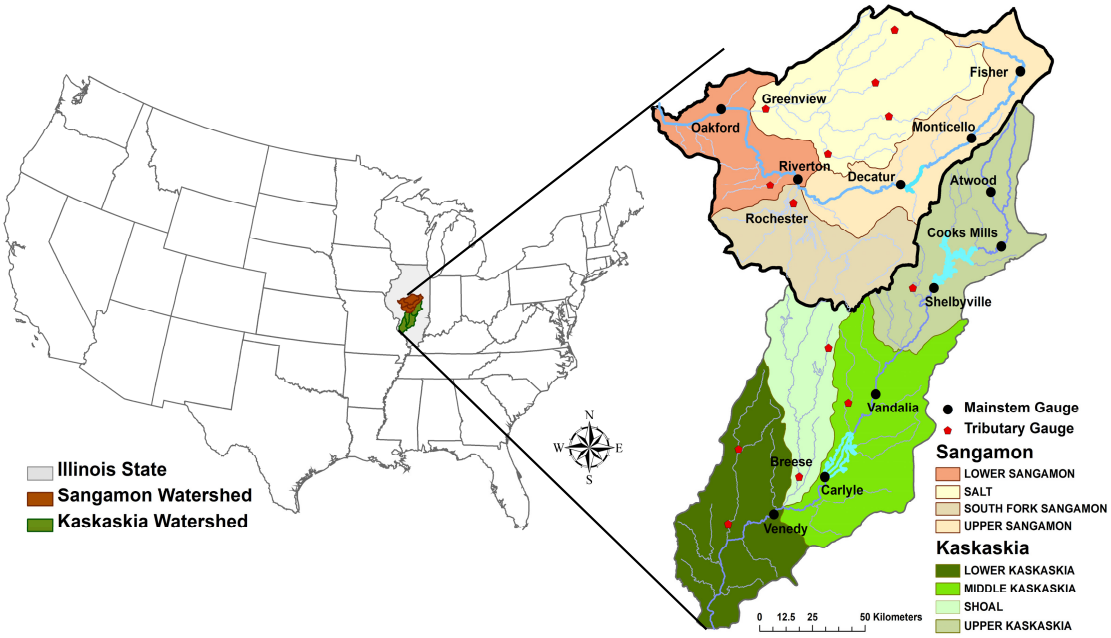

Figure 3.2: Locator map for the Sangamon and Kaskaskia watersheds with gauge site and stream details.

3.3.2 Study area – watershed spatial characteristics

The study area consists of two watersheds in central IL (Figure 3.2). To the north is the 14,000 km^2 Sangamon River watershed, a tributary of the Illinois River which itself is a tributary of the Mississippi River. To the south is the Kaskaskia River watershed, a roughly 15,000 km^2 tributary of the Mississippi River. Although each watershed consists of four main hydrologic units (HUCs), the hydrologic connectivity differs between them. The Sangamon watershed consists of an Upper and Lower main-stem and two hydrologically separate tributaries – Salt Creek in the north and South Fork in the south, with confluences in the Lower Sangamon. The Kaskaskia watershed consists of an Upper, Middle, and Lower main-stem and one large tributary, Shoal Creek. There are three in-line reservoirs of note in these watersheds, each with a different purpose and release rules. In Sangamon, Lake Decatur has provided water supply for the city of Decatur since its construction in 1922; in Kaskaskia, the more recent Lake Shelbyville in the north mainly provides recreation, while Lake Carlyle in the south provides water supply, recreation, and flood control for both the Kaskaskia and Mississippi Rivers. Locations of these reservoirs as well as the main-stem and tributary stream gauge locations are shown in Figure 3.2. On average, the region receives about 1000 mm of rain annually, although there can be a great deal of spatial variability due to a higher frequency of high-intensity convective rainfall events that cover localized areas compared to less-intense frontal events that cover large areas. Precipitation is slightly seasonal here, with the spring months being wetter on average than the rest of the year, followed by drier late summer months that closely follow the period of peak evaporative demand.

The legacy of the geologic history of these two watersheds manifests in their topography and soils, as well as the vegetation that has co-evolved with both, under influence of the climate. The relatively flat topography (Figure 3.3a) is a result of glaciation, and in the case of the Sangamon watershed, repeated glaciation. There is more topographical relief in Kaskaskia compared to Sangamon; in Sangamon, a gentle ridge divides the watershed about midway downstream, separating an extensive, flat upland region from the mostly flat lowland region. The same ridge that divides the Sangamon watershed extends into

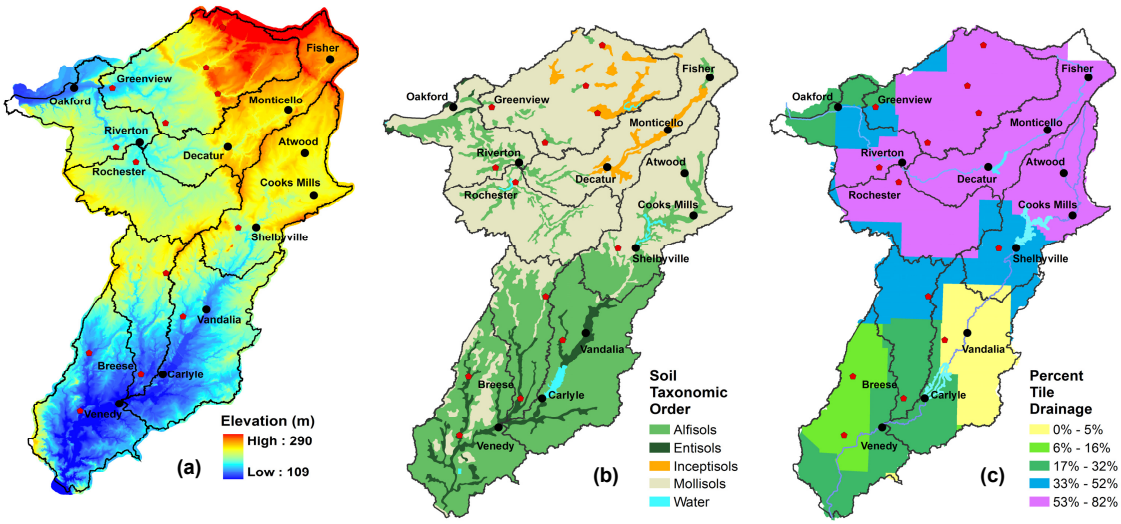

Figure 3.3: USGS topography (a), STATSGO soil taxonomy (b), and WRI estimates of percent tile drainage by county (c).

Kaskaskia, separating the upper portion, which shares soil, vegetation, and drainage characteristics with Sangamon, from the rest of the Kaskaskia watershed, which is marked by gentle slopes to deeper valley bottoms. These valley slopes provide enough topographical relief that the Middle and Lower Kaskaskia were more well-drained naturally (Prince, 2008) than the Sangamon and the Upper Kaskaskia.

The geologic history also had significant implications on the soils and vegetation that co-evolved in these two watersheds: to the north, more recently glaciated portions of both watersheds developed mollisols rich in organic matter under dense grassland, while to the south, weathering produced alfisols that developed mainly under forest vegetation (Figure 3.3b). The pre-settlement land cover (Figure 3.1a) illustrates this well – Sangamon and Upper Kaskaskia were dominated by prairie grassland vegetation, while the rest of Kaskaskia was mostly forest vegetation. Estimating from the IL Natural History Survey dataset, the Sangamon was roughly 90% prairie while the Kaskaskia was about 85% forest. It is inter-

esting to note that present-day land cover in Sangamon and Upper Kaskaskia still follows these proportions, although the prairie grasses have been replaced for the most part by row crop agriculture. Even though much of the forest in Kaskaskia is gone, pasture and forest vegetation together comprise about 30% of the watershed, compared to only 10% in Sangamon.

Both the Sangamon and Upper Kaskaskia require extensive tile drainage (Figure 3.3c) for crops to be successfully grown, due in part to a hardpan subsoil layer combined with lack of a natural surface drainage network resulting from their flat, upland topography. In the southern portions of the Kaskaskia watershed, installation of tile drainage is impeded by a shallow fragipan layer; however, this is not a problem for agriculture since waterlogging is less common here, due in large part to the higher slopes the landscape here tends to have and which may also explain why deep-rooted forest vegetation was able to become established here.

3.3.3 Streamflow analyses

For the comparative analysis of streamflow time series data, observed daily streamflow data for 24 USGS gauges, 12 each in both Sangamon and Kaskaskia watersheds, was obtained for a 22 year period from 1990 through 2011. This time period was chosen to reflect the current condition of the watersheds, which would include the cumulative effects of past land use/ land cover and climate changes. Unfortunately, very few gauges have sufficient historical data to perform the same analyses for the 1940s or earlier as a comparison. Of the 12 gauges chosen for each watershed, 5 were located on the mainstem of Sangamon and 6 on the mainstem of Kaskaskia, with the remaining gauge locations distributed among the tributary streams in each watershed (Figure 3.2). Gauges were chosen primarily for length of record; there are many more gauges in these watersheds that have been installed recently, but their period of record is too short for meaningful analysis. To aid comparison between sub-watersheds of differing sizes, the daily instantaneous flow rate (ft^3/s) from the gauge records was converted to daily volume ($m^3 day^{-1}$) and then scaled by drainage area (m^2) and converted to mm to give water yield.

Hydrologic signatures at three timescales were then calculated from the daily record. The inter-annual variability signature sorts annual total streamflow from highest to lowest exceedance, with a flat slope to the plot signifying low year-to-year variability in streamflow. The regime curve (RC) plots the monthly average streamflow over the year and shows the seasonal streamflow patterns within the year. Lastly, the flow duration curve (FDC) plots daily streamflow magnitude (on a log scale) as a function of the percent of time it is exceeded. The FDC can be divided into 3 sections, with the upper third corresponding to flood events and the fast flow response of a catchment, the middle third to the slow flow response of the catchment and its regime curve, and the lower third the low flow response (Yokoo and Sivapalan, 2011). While the upper third of the FDC is mainly related to precipitation intensity, the lower third is more related to properties of the catchment itself (Yaeger et al., 2012). The lowest streamflows occur in the late summer to early fall when evaporative demand still exceeds precipitation but when crop vegetation (although not forest vegetation) is beginning to die off. With the proposed land use changes affecting the length of the growing season and thus the amount of ET during this time, the present study will specifically focus on the low flows.

Because the historical analysis of the region indicated that baseflow had increased over time due to changes in crop rotations and expanded tile drainage, the baseflow index (BI), defined for a given time period as the ratio of mean annual slow flow to mean annual total flow, was also calculated. Separation of daily streamflow at each gauge into the fast and slow components was accomplished with a simple one-parameter low-pass filter as described in Ye et al. (2012).

Examination of the monthly regime curve showed that, on average, summer flows are smaller than winter flows; thus only the summer low flow period, defined here as the days of July 15 through November 15 of each year in the period of record, was chosen for the low flow analysis. The streamflow threshold method was chosen to determine streamflow drought periods (Yevjevich, 1967; Zelenhasić and Salvai, 1987)); at each gauge the area-averaged streamflow (water yield) corresponding to both the 90th and 95th percentiles on the flow duration curve was used as the threshold (Zelenhasić and Salvai, 1987). After the initial analysis, it was determined that the 95th percentile flows in this region were too

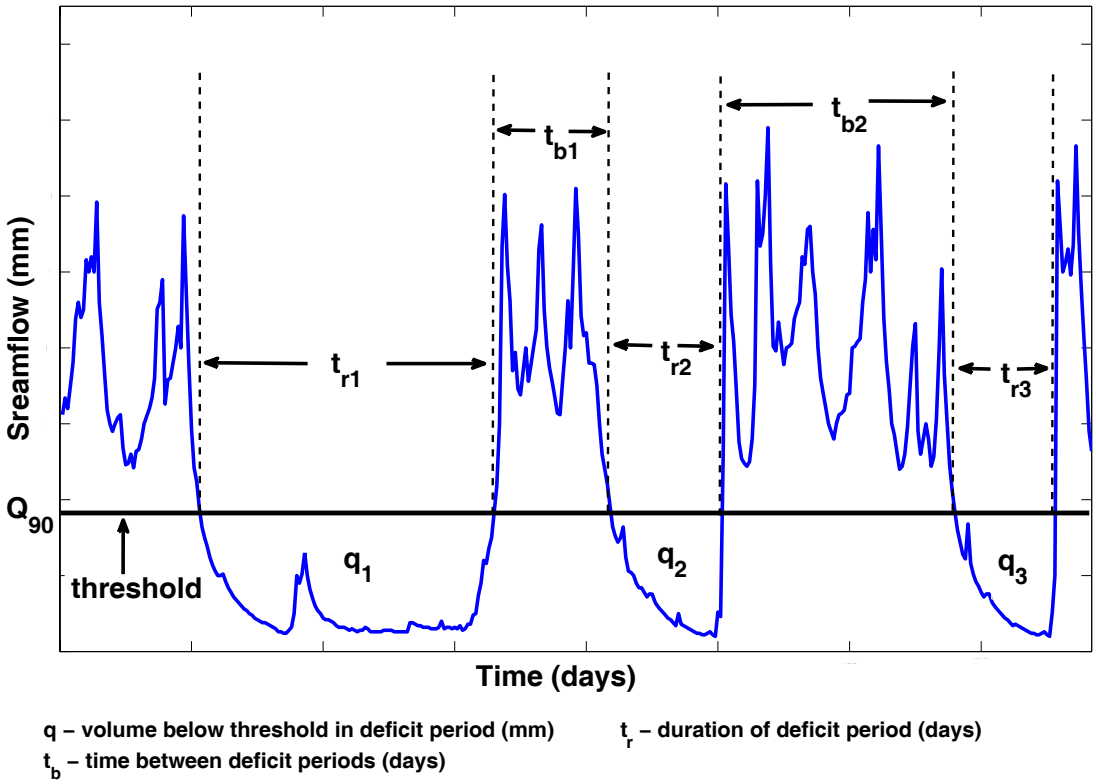

Figure 3.4: Definition of deficit, deficit period, and occurrence interval used in low flow analysis.

small to produce meaningful results; thereafter, only the flow corresponding 90th percentile (Q_{90}) was used, and these results are presented later in this paper. A streamflow deficit was determined to occur if the daily flow (Q_d) during the summer low flow period was less than the threshold flow (Figure 3.4).Since daily area-averaged flow was used, the total deficit volume for a given deficit period can be defined as

$$q = \sum_{d=1}^{t_r} (Q_{90} - Q_d) \tag{3.1}$$

where $d = 1$ is the first day where $Q_d < Q_{90}$, t_r is the total duration (days), and Q is the total volume (mm) for that duration period. If there was more than one deficit period in a low flow season, the time between deficit periods t_b (days) was also recorded; if only one deficit period occurred in a year, t_b was recorded as 0. This analysis was repeated for each year in the period of record. In some cases, a day or two where $Q_d > Q_{90}$ separated two

longer deficit periods. According to Zelenhasić and Salvai (1987), these two deficit periods are likely to be related, and thus should be considered one long deficit period. To prevent related deficit periods from being counted as separate events, a threshold t_b of 3 days was chosen based on examination of the low flow behavior of a few gauges so that in a given year, consecutive deficit periods having fewer than 3 days between them were merged into one deficit period. Lastly, the mean and standard deviation were calculated for Q, t_r, and t_b to give an overall picture of the low flow response from these watersheds. Even after normalizing by drainage area, there was sufficient variability in the threshold flows that the average relative deficits (RD), defined here as the ratio of the average deficits to their respective Q_{90}, was also calculated (Hisdal and Tallaksen, 2003).

Lastly, to gain a quantitative idea of the spatial and temporal variability of precipitation in this region, 9 points were selected covering the spatial extents of the two watersheds; from west to east, 3 points were chosen across the northernmost latitude, 3 across the southern, and 3 across the divide between the two watersheds. Annual and monthly precipitation data was obtained from the PRISM dataset Table 3.1 at each of these locations and was used to calculate both the inter-annual variability and monthly regime curves for the precipitation.

3.4 Results

3.4.1 Data description

The inter-annual variability signature of the streamflow of these watersheds (Figure 3.6) shows that there is some variability in total streamflow between years, especially for the driest years, but less so in Kaskaskia. The differences in annual streamflow between the two watersheds can be explained in part by the spatial variability in precipitation described in Section 3.3.2 and which is shown in Figure 3.5, as well as the catchment response to it. Slower-draining catchments, such as those in Middle and Lower Kaskaskia, tend to even out this spatial variability in precipitation, but the extensive tile drainage in the Sangamon mitigates this filtering effect. In addition to this localized variability, there is also a regional north-south gradient to the annual precipitation, with the southern portions of Kaskaskia

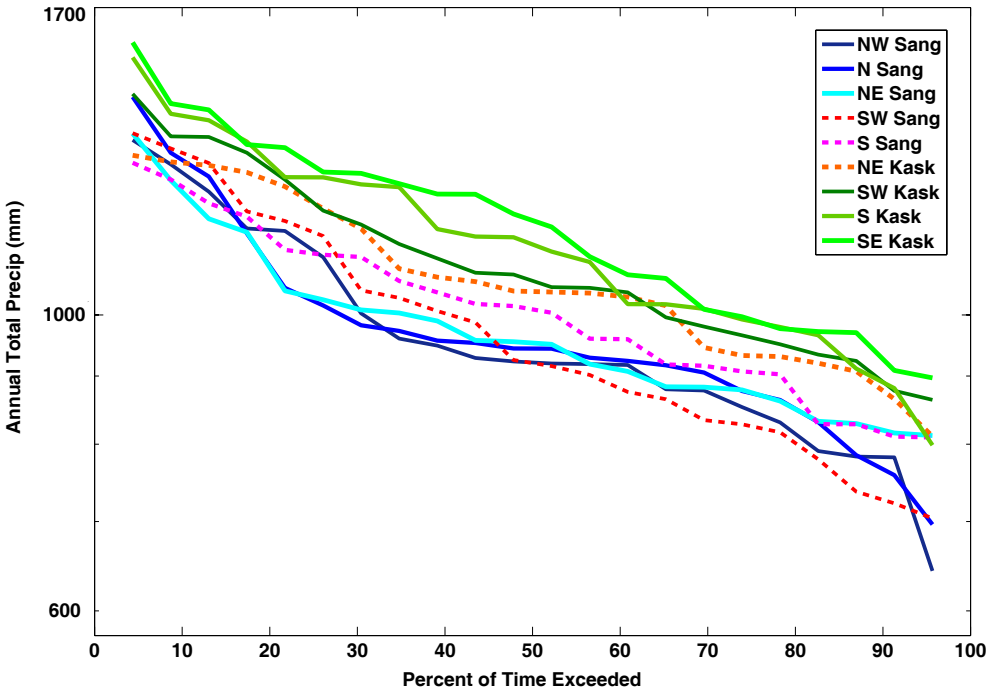

Figure 3.5: Precipitation inter-annual variability for both watersheds for 1990 – 2011.

on average receiving more rainfall annually than the northern parts of Sangamon (Figure 3.5). Of interest is Sugar Creek in the Sangamon (Figure 3.6, light blue dashed line), a small (89 km^2) headwater stream in the upper Salt Creek HUC. Why such a small stream experiences such high runoff is not explained by climate variability alone. This represents another human impact signature: the local municipality discharges treated wastewater, the source of which is likely groundwater or water supply from a lake outside the watershed, into Sugar Creek, thus artificially inflating the hydrograph.

The regional variability in precipitation also manifests at the monthly scale (Figure 3.7) with the north-south gradient evident in the winter and early spring, but not so in summer and early autumn. For the seasonal pattern of monthly average streamflow, aside from Sugar Creek, there is very little variation between sub-catchments in the Sangamon watershed, even immediately downstream of Lake Decatur (Figure 3.8). Winter flows are larger on average than summer, and the flood peaks all occur in May, coinciding with peak rainfall for the most part and the very beginning of the growing season. There is a smaller, secondary peak in March, most likely due to spring rainfall and snowmelt. From May to June there is only a

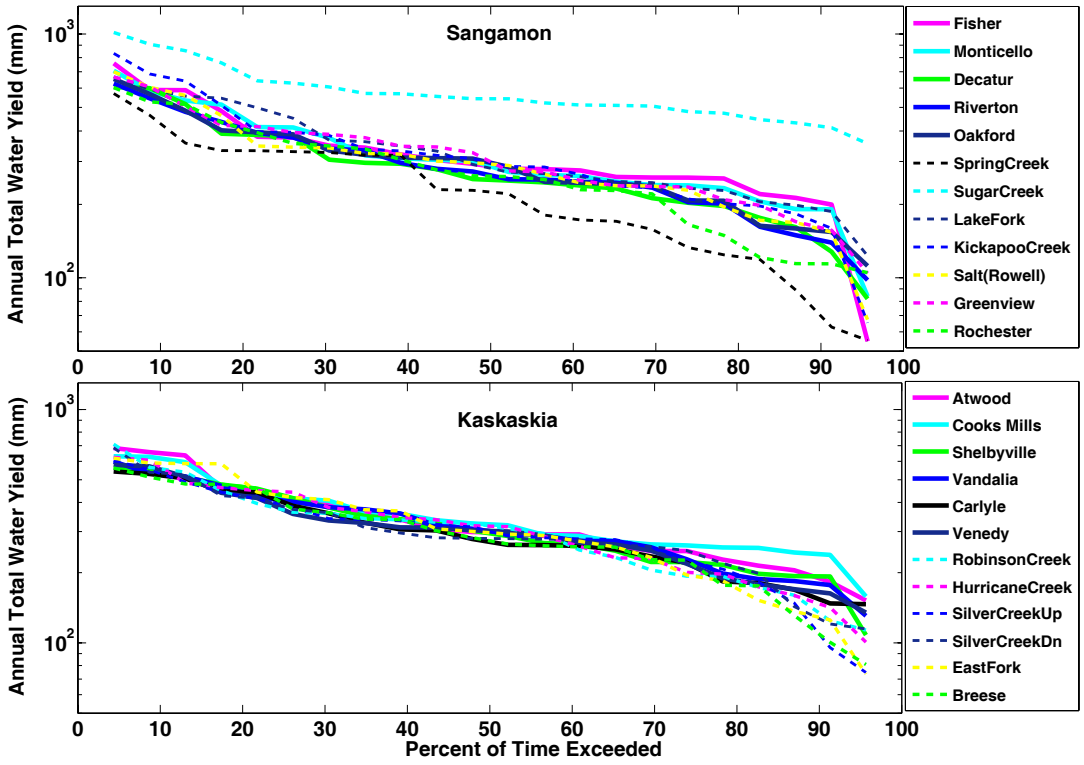

Figure 3.6: Streamflow inter-annual variability for both watersheds for 1990 – 2011. (Note: main-stem gauges are shown as solid lines and tributary gauges are shown as dashed lines).

slight decline in monthly flow; this is because the crops are not yet at peak water usage and the tile drains move excess infiltration (once it reaches them) quickly to drainage ditches and streams. The steepest decline happens later from June to July, when crop ET is at its peak. In contrast, there is much more variability in the seasonal streamflow patterns for Kaskaskia. The influence of the two reservoirs in the Kaskaskia watershed can clearly be seen in the Shelbyville, Vandalia, Carlyle, and Venedy regime curves: winter flow is increased (to lower the water level in the reservoir), and the May flood peak is greatly reduced, with the excess water slowly released so that the lowest monthly flow (for this time period) is shifted from August to September and October. The flow regimes for the rest of the Kaskaskia gauges, located on uncontrolled reaches, show a general seasonal pattern similar to that in Sangamon, although a closer look reveals an interesting exception. The regime curves for Kaskaskia tributary streams (Figure 3.8, dashed lines) show a sharp, almost linear decline from May to July, while those for Atwood and Cooks Mills, two main-stem gauges located

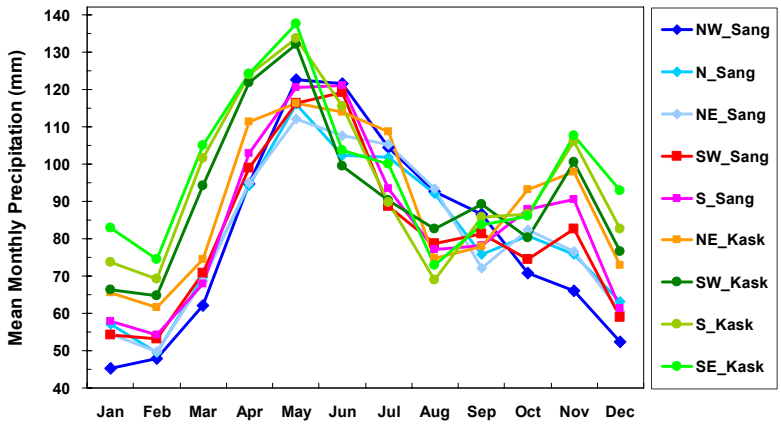

Figure 3.7: Monthly precipitation regime curves for both watersheds for 1990 – 2011.

above Lake Shelbyville in an area where land use, soil type, and tile drainage are similar to those in Sangamon, show the same small decline from May to June and sharp decline from June to July that was seen in Sangamon RCs. Thus, three distinct flow regime patterns are discernible in Kaskaskia: reservoir-dominated, row-cropped and tiled-drained, and less artificially drained with more heterogeneous land cover.

For the monthly RC, the streamflow patterns for the Sangamon were much less variable than those for Kaskaskia. For the daily FDC, however, the reverse is true, especially for the low flows (Figure 3.9). There is more variability in the lower tail (66% – 100% exceedance) in the Sangamon Basin compared to the Kaskaskia. Another human impact on streamflow can be seen at the Decatur gauge. Lake Decatur supplies water to both the City of Decatur and a large ethanol refinery nearby. The Decatur gauge shows no flow almost 5% of the time, while upstream (Monticello) and downstream (Riverton) do not. This is likely due to the demand plus evaporation in summer being greater than the inflow from upstream. Downstream of the Decatur gauge is the wastewater discharge point for the city; thus the Riverton gauge further downstream does not manifest the expected decrease in flow. The middle limb of the FDC, which represents the average, regime flows at each gauge, shows little variability within the Sangamon watershed, just as was seen in the RCs (Figure 3.8).

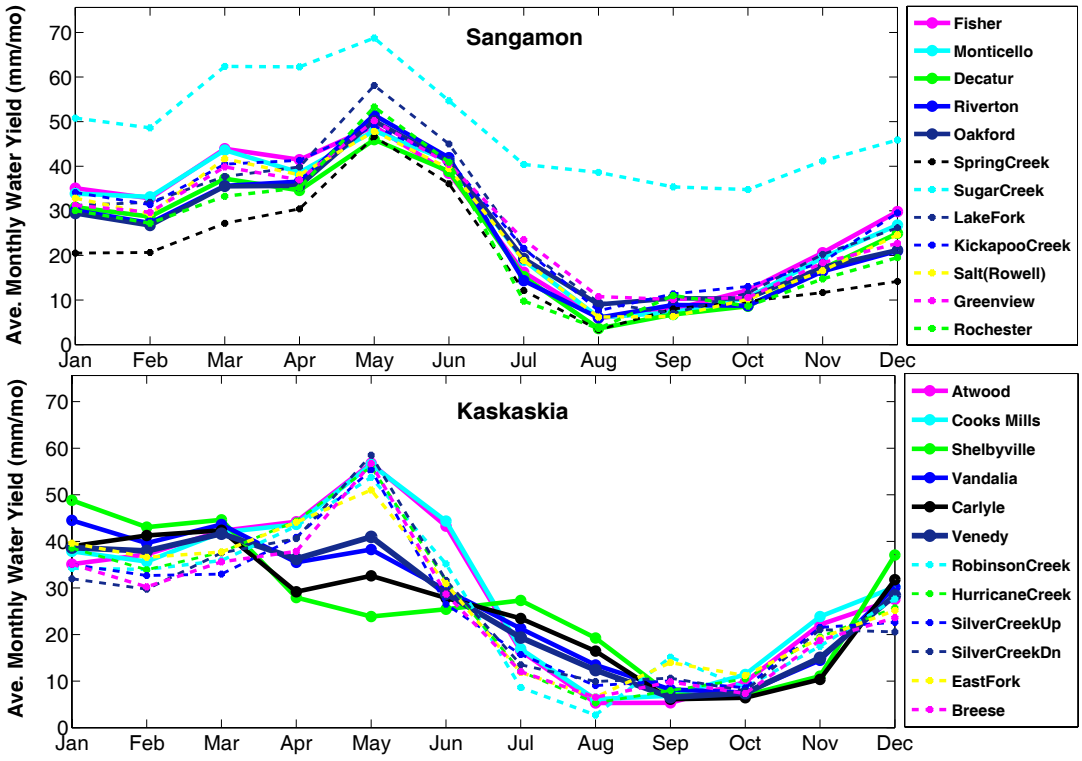

Figure 3.8: Monthly regime curves for both watersheds for 1990 – 2011. (Note: Main-stem gauges are shown as solid lines and tributary gauges are shown as dashed lines).

Also as seen in the RCs, the reservoirs' influence is clearly evident in the Kaskaskia FDCs as increased variability in the upper portion of the FDC (high flows reduced on controlled reaches) and convex curvature to the middle limb of the FDC (increasing mean flows of the same reaches). Of note is the much smaller variability in the low flows in Kaskaskia compared to Sangamon, which could be due to there being fewer drainage tiles in Kaskaskia. In that sense, streamflow in Kaskaskia would then be more "filtered" than streamflow in Sangamon; without tile drainage, inputs are retained longer and thus the output is more processed by the catchment. It is also interesting to observe that the gauges downstream of Sugar Creek (Greenview and Oakford) have consistently large low flows, even at 95% exceedance; this is likely due to the effects of the artificially-high flows in Sugar Creek propagating downstream. The flat slope of some of the low flow tails of the Sangamon FDCs would also seem to indicate increased baseflow compared to Kaskaskia.

In general, as drainage area and streamflow increases, heterogeneity decreases, and this

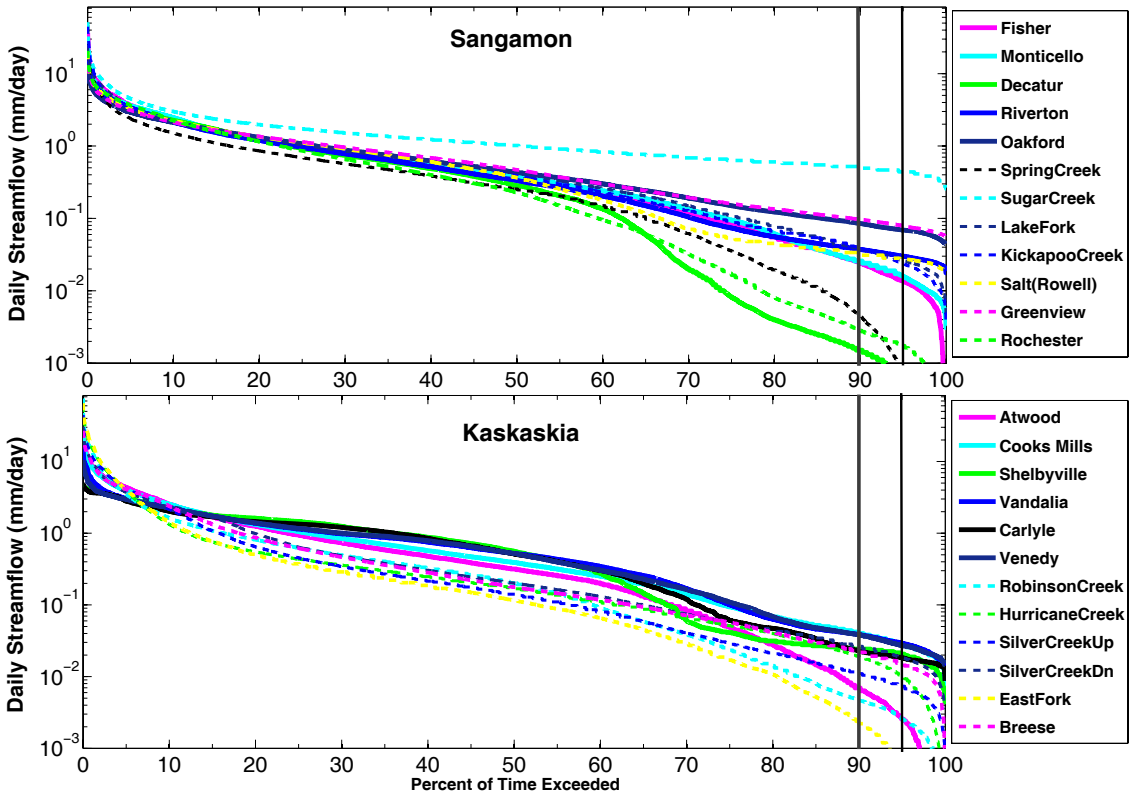

Figure 3.9: Flow duration curves of both watersheds for the period 1990 – 2011. (Note: Main-stem gauges are shown as solid lines and tributary gauges are shown as dashed lines).

pattern can be seen, for Kaskaskia at least, even for the lower portion of the FDC (here the 65% - 95% exceedance). When low flow FDC quantiles of daily water yield for both watersheds are plotted against their respective drainage areas, there is less of a trend in the Sangamon quantiles compared to Kaskaskia (Figure 3.10). An interesting outlier in this figure, shown in the circled portion of Figure 3.10a, is the Sugar Creek catchment, where streamflow has been artificially increased due to discharge of treated wastewater. Since there is less variability in human water use, overall streamflow variability decreases; this was also seen previously as a much flatter FDC compared to other streams in the watershed (Figure 3.9). Furthermore, based on regression analysis for Q_{65} (Figure 3.10, brown lines) and Q_{95} (Figure 3.10, blue lines), we expect the magnitude of water yield to increase with an increase in drainage area; however, two mid-size sub-watersheds (Decatur and Rochester) have water yield magnitudes similar to a sub-watershed an order of magnitude smaller in size (Spring Creek, in Lower Sangamon). From the FDC, we see that all three of these sub-watersheds

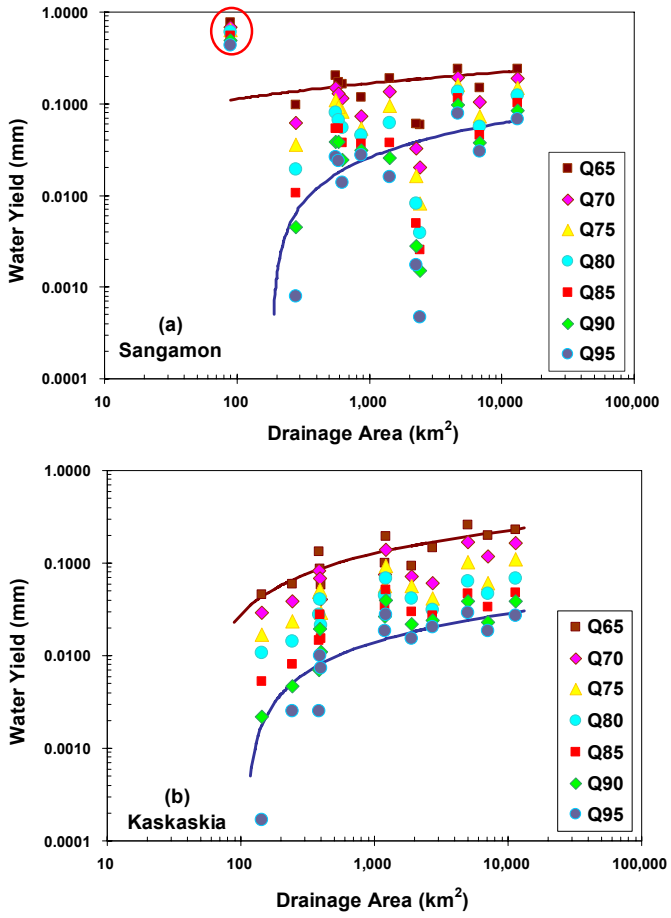

Figure 3.10: Scaling of the lower limb of the FDC with drainage area in (a) Sangamon and (b) Kaskaskia. Note: the circled area in (a) is the Sugar Creek catchment.

occasionally experience zero daily flow (Figure 3.9). For Spring Creek, this is likely due to the small drainage area (<100 km^2); for Decatur and Rochester, this is more likely due to municipal and industrial water supply diversions that bypass the flow gage, but the diverted water remains within the Sangamon Basin and thus does not impact downstream gages. In Kaskaskia, however, there appears to be a more structured, nonlinear relationship with drainage area for the smallest flows, becoming more linear as drainage area increases. For both watersheds, smaller drainage areas with extensive tile drainage tended to experience lower extreme low flows than did non-tile drained areas; as drainage area increases this effect becomes less clear due to an increasing combination of effects of other catchment modifications.

In summary, there is some variability in streamflow, even at the annual scale, most noticeable in dry years, and likely a catchment response to spatial variability in precipitation. At the monthly scale, the high level of tile drainage and similarity in land use, and lack of controlled reservoirs in the Sangamon watershed results in most of the gauges responding in a similar fashion throughout the year. Kaskaskia on the other hand, shows a much more heterogeneous seasonal response, due to the presence of two large controlled dams, a relative lack of tile drainage, and more heterogeneous land cover. Finally, at the daily scale, the catchments in Kaskaskia produce much more uniform FDC low flow tails than do the catchments in Sangamon. Although the FDC has limitations – the linear time element has been removed – the higher variability seen in the low flow tails has important implications for the proposed biofuel crop changes, as we expect that the major impacts of this land use change will be felt here. To understand what these impacts might be, we begin with a low flow analysis of the late summer and early fall streamflow period.

3.4.2 Low flow analysis

Overall, average deficit volumes in the Kaskaskia watershed are fairly small (<0.2 mm), with an average time between, in most places, of about a week, but varied in their average duration. However, some of the threshold flows were also very small, particularly those of headwater streams; as a consequence of the very small denominator, Kaskaskia showed the largest average RD of either watershed (10 – 13 times the Q_{90} threshold). For the rest of the watershed, the average RD of the mainstem reaches was, on average, 2.5 to 4.5 times the threshold flow, while the RD of tributary reaches was, with the exception of the southernmost gauge, a little higher (Figure 3.11a). Like Kaskaskia, the headwater mainstem reach in Sangamon showed a large average RD (8 times the threshold), but for the rest of the Sangamon watershed there does not appear to be a strong spatial pattern to the average RD as there was in Kaskaskia. Here, some tributary reaches experienced larger average RD (7.0 – 9.5 times than the threshold) while others did not, and aside from Riverton, the Sangamon main-stem RD were larger than those in Kaskaskia. There was also more spatial variability in the average magnitudes of t_r and t_b in the Sangamon watershed compared to Kaskaskia;

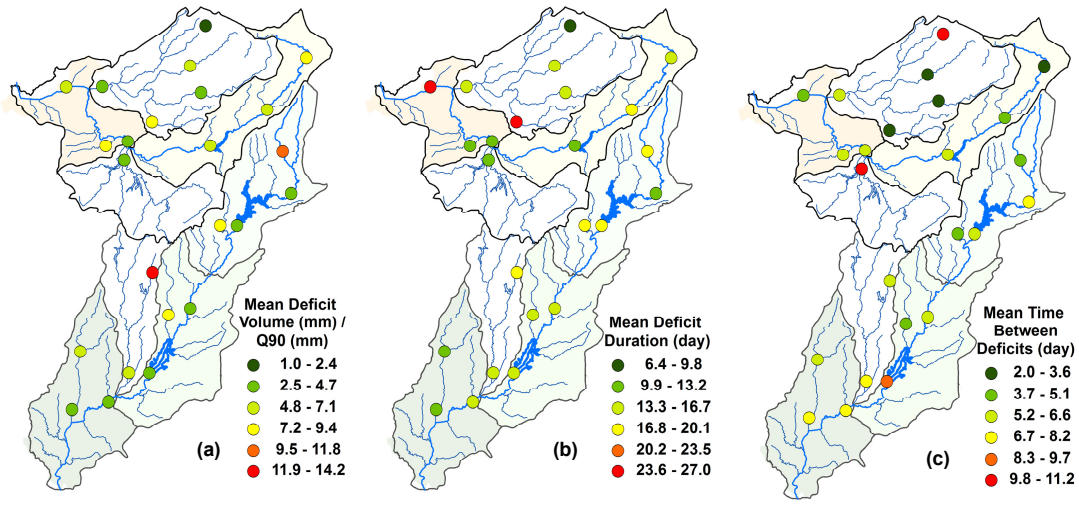

Figure 3.11: Relative deficit (a), mean deficit duration (b), and average time between deficits (c) for both watersheds.

in Sangamon, mean deficit durations ranged from the highest to the lowest values (6 – 30 days), while in Kaskaskia the range was much smaller (10 – 18 days) (Figure 3.11b). The same was true for the time between deficit periods; in Kaskaskia, t_b varied between 4 to 8 days, while in Sangamon the range was 2 to 11 days (Figure 3.11c). For the study period, on average, deficit periods were shorter and more frequent in Sangamon than in Kaskaskia (Figure 3.12).

Figure 3.13 presents the low flow analysis in terms of scaling with drainage area. Again, there is more scatter in Sangamon compared to Kaskaskia, although both watersheds show a general trend of decreasing average RD with increasing drainage area. This pattern of higher variability for Sangamon compared to Kaskaskia also held for mean deficit duration and time between deficits when plotted against drainage area (not shown for sake of brevity). It should be noted that the one outlier in Figure 3.13 is the Sugar Creek gauge, where wastewater discharge from sources outside the basin have artificially increased the streamflow.

The baseflow analysis yielded interesting results. Overall, BI for the Sangamon gauges

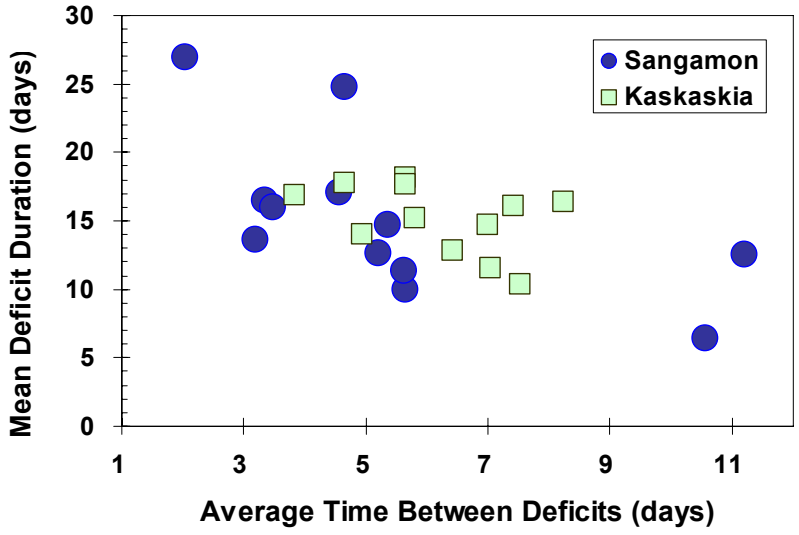

Figure 3.12: Average frequency vs. duration of deficit periods from 1990 – 2011.

ranged from 0.5 to 0.69, with the exception of two downstream gauges, Greenview at the outlet of the Salt tributary and Oakford, the far downstream main-stem gauge, both of which are downstream of the Sugar Creek gauge as well as being channelized and straightened. Three different groupings can be seen in the Kaskaskia BI values (Table 3.2), which correspond to the three flow regime patterns seen in the regime curves (Figure 3.6. The less tile drained sub-watersheds in Kaskaskia with more heterogeneous land cover showed BI <0.5, while the two gauges in the tile-drained, row-cropped Upper Kaskaskia region had BI values similar to those in Sangamon. All main-stem gauges below the reservoirs in Kaskaskia, as well as the two downstream gauges in Sangamon mentioned above, showed BI >0.7. Channelization of river reaches increases baseflow by increasing the area in contact with deeper groundwater sources; flood control reservoirs increase baseflow by releasing excess storage during natural low flow periods in order to ensure sufficient storage for the high flow periods. Lake Decatur is not a flood control reservoir, and so does not show this release pattern.

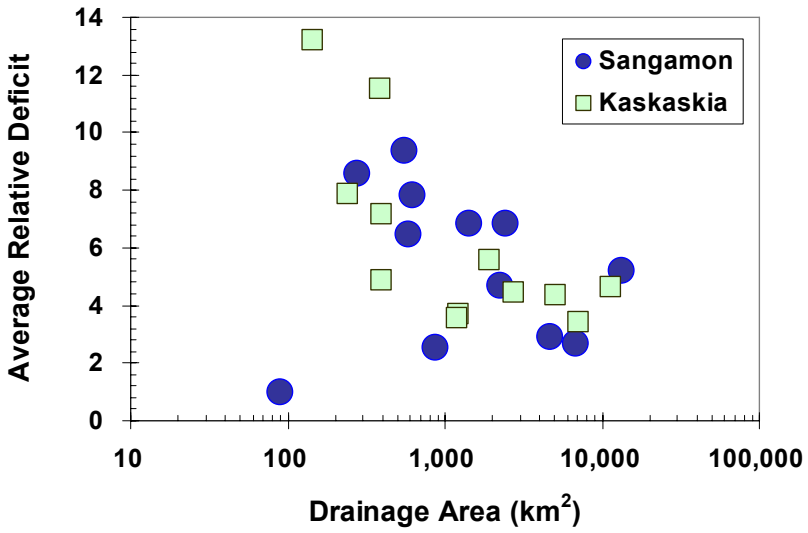

Figure 3.13: Scaling of relative deficit with drainage area.

3.5 Discussion

3.5.1 Legacies of the past manifested in current hydrologic responses

The case study catchments reflect in many ways the legacies of their pasts. The glacial histories of each watershed influenced the soils and vegetation that coevolved, and these, combined with the topographical differences between the two watersheds influenced the amount and type of drainage European settlers needed to implement to prepare the land for agriculture. In Sangamon and Upper Kaskaskia, a shallow groundwater table combined with a flat topography required extensive surface and subsurface drainage to prevent waterlogging and allow crops to flourish. In the past, this region was predominately wet prairie; today it is predominately annual row crop agriculture. This has manifested in the catchment response as a homogeneous seasonal flow regime, increased variability in the low flows, and an higher BI. The rest of the Kaskaskia watershed has very little tile drainage for a combination of reasons, all of which are linked to historical impacts. The glacial legacy of this region resulted in greater topographical relief so that it is more well-drained naturally, and flatter, waterlogged portions are sufficiently improved with only surface drainage modifications.

Table 3.2: Baseflow Index values for 1990 – 2011 and associated catchment features for (a) Sangamon and (b) Kaskaskia.

(a) Sangamon	**BI**	**Feature**	**(b) Kaskaskia**	**BI**	**Feature**
Fisher	0.58		Atwood	0.56	Tiled,
Monticello	0.61		CooksMills	0.59	row cropped
Decatur	0.53		Shelbyville	0.75	
Riverton	0.63		Vandalia	0.71	Reservoirs or
SpringCreek	0.58	Tiled,	Carlyle	0.78	channelized
SugarCreek	0.65	row cropped	Venedy	0.75	
LakeFork	0.63		RobinsonCreek	0.46	
KickapooCreek	0.60		HurricaneCreek	0.33	Less tiled
Rochester	0.59		SilverCreekUp	0.36	row cropped
Salt_Rowell	0.67		SilverCreekDn	0.44	pasture
Oakford	0.71	Reservoirs or	EastFork	0.30	forest
Greenview	0.70	channelized	Breese	0.47	

Furthermore, the shallow fragipan layer underlying much of the watershed makes extensive tile drainage infeasible. Historically, this region was predominately forested, and although the present-day land cover is predominately annual row crop agriculture, there is also a large amount of pasture land and forest as well. In addition, large flood control reservoirs constructed in Kaskaskia to regulate streamflow also provide a water quality benefit; because of the increased residence times, the total mass output of nitrate-N (NO_3-N) from Kaskaskia is less than that from Sangamon. These differences in land cover, subsurface drainage, and reservoir operation have all manifested in the catchment response as three distinct groupings in the seasonal flow regime, and these patterns are also reflected in the BI.

The overall effect of tile drainage on the hydrologic response is one of introducing a new threshold. Infiltrated precipitation reaches the low-permeability sub-soils and begins to saturate the soil column until it reaches the depth of the tile drains, which then provide a preferential flow pathway toward the surface drainage network. This reduces the time that the soil layer above the tile drain is waterlogged and thus damaging to crop roots. It also limits replenishment of root zone soil moisture from the saturated layers below. Thus,

in the summer growing season, when there is less precipitation and evaporative demand is greatest, the major source of soil moisture is from through-fall. During this period, smaller ditches and streams dry up as the soil does, while larger streams are fed by baseflow from deeper soil moisture storage. This can be seen in the summer low flow period in the RC (Figure 3.8), and in the FDC (Figure 3.9) of smaller tributary streams where flow is not exceeded 100% of the time. The low flow analysis revealed that on average, the Sangamon watershed experienced shorter but more frequent streamflow deficit periods than did Kaskaskia during 1990-2011. However, there was also more variability associated with deficit periods in Sangamon, compared to Kaskaskia, due in part to drainage effects on the catchments' response to localized spatial variability in precipitation. A recent study of hydrologic drought by Van Lanen et al. (2013) found that groundwater, as manifest in catchment response times, is a more important control on drought duration and deficit magnitude than either climate or soil type, and that fast catchment response times increased the frequency of deficit periods. All these findings could have important implications for the sustainable large-scale planting of cellulosic biofuel energy crops such as *Miscanthus* in this region.

3.5.2 Implications for biofuels land use change

As mentioned previously in Section 3.1, *Miscanthus* is a perennial grass, which means it is not planted according to an annual timetable, as corn and soybeans are, but emerges from winter dormancy when soil and air temperatures are sufficiently warm. This typically occurs around the beginning of April, as opposed to corn and soybeans which are usually planted starting in mid April and starting in early May, respectively. At the end of the growing season, ET of corn and soybeans declines and ends a few weeks before harvest, which usually occurs at the beginning of October, but *Miscanthus* continues to transpire until the first frosts, usually in November (Schilling et al., 2008; McIsaac et al., 2010; Le et al., 2011), so that its growing season is two months or more longer than the current row crop annuals. In addition, the canopy of *Miscanthus* is much more dense than that of corn or soybeans, with a maximum leaf area index (LAI) of up to 10 compared to 7 for soybeans

and 5.5 for corn (Heaton et al., 2008; Le et al., 2011) which would also intercept more water than the annual crops, thus reducing through-fall in addition to drying the soil profile earlier in the year (McIsaac et al., 2010). In a tile drained watershed, this could mean that the local streams become baseflow-dependent earlier in the year than under row crops, and, in dry years especially, experience larger and/or longer deficit periods during the summer low flow season. On the other hand, the greatly reduced fertilizer requirements of *Miscanthus* would mean a significant decrease in nitrate export from tile drained watersheds.

The increased fine scale heterogeneity brought to light by the analysis of recent streamflow data is a complicating factor when predicting how these catchments would respond to large-scale planting of *Miscanthus*, since knowledge from one location within the watershed cannot be directly transferred to another, similar location in the same watershed. When combined with the inherent spatial variability of the precipitation inputs, we see there are limitations to what can be learned from observed data. However, much modeling work has been done to try to predict the impacts of this new crop on the hydrology of the region, both at the watershed and at river basin scales. Using a dynamic global vegetation model to simulate crop ET in the Midwestern US, Vanloocke et al. (2010) found statistically significant increases in annual ET and a corresponding decrease in hillslope drainage even for moderate fractions of land planted with *Miscanthus*. The impacts of this change, however, varied with location; where water resources were already stressed, large-scale planting of *Miscanthus* tended to worsen conditions. More recently, a smaller scale study of the Iowa River basin (Wu and Liu, 2012) using the watershed hydrologic model SWAT obtained similar hydrologic results. A second large scale study used the SWAT model to examine the effects of planting Switchgrass, another proposed biofuels crop, in the Upper Mississippi River basin (Wu et al., 2012). Similar to what has been found for *Miscanthus*, converting land to switchgrass reduced nitrogen export and soil erosion but also increased ET and decreased the baseflow component of streamflow. Each of these studies covered similar regions of the US where there is extensive tile drainage, and all shared a common result that an increase in fraction of land covered by *Miscanthus* led to an increase in annual ET and a corresponding decrease in hill slope drainage to streams. The hydrologic impacts, however, varied with the scale of the study, the fraction of *Miscanthus* investigated, the crop that was replaced by *Miscanthus*, and, for

the larger-scale studies, the location within the watershed as well.

The threshold effect of the tile drains coupled with the switch from perennial and sod vegetation to intensive annual row crop agriculture led to an overall decrease in ET and an increase in baseflow. Large-scale planting of *Miscanthus*, a perennial grass, would have the opposite effect, dependent on the location and amount planted, and current hydrologic conditions, creating a kind of biofuels threshold effect. As the fraction of *Miscanthus* in a watershed increases toward this threshold of sustainability the effects will likely be seen in the low flow characteristics, with a corresponding increase in streamflow deficits and deficit durations and a decrease in the time between deficit periods. These effects would also vary spatially within and between watersheds. We have seen that small, tile-drained subcatchments in Sangamon and Kaskaskia already experience smaller low flows than those without tile drainage. Because low flows in this region are dependent on shallow groundwater, such subcatchments would be more sensitive to the reduction in hillslope drainage that would result from planting *Miscanthus*. In tile-drained watersheds with surface water demands in upstream regions, such as Sangamon, this could reduce inflow to important water supply reservoirs; in Kaskaskia, however, these effects may be less noticeable because of the relative lack of tile drainage in the majority of headwater subcatchments.

3.5.3 Study limitations and future work

One limitation to this type of data-based analysis is that there are often too few gauges with sufficiently long periods of record upon which to base such analyses, thus creating "gaps" in the knowledge. This is where the large-scale watershed model becomes useful; once the basin is delineated into smaller sub-catchments, modeled streamflow information is then available at the outlet of each one, thereby greatly increasing the spatial resolution of the catchment response and allowing a more complete picture to emerge. Furthermore, the empirical analysis of these case study watersheds revealed that the past is not necessarily a deterministic predictor of the future because of the heterogeneous response created by the extensive drainage modifications to the system. However, understanding gained from such analysis can be used to inform the modeling process; thus both empirical and physically-

based modeling analyses are needed to make informed predictions about the possible effects of proposed land use changes. The next step, therefore, is to model this replacement of current row crop rotations with *Miscanthus* or other biomass energy crops.

The study region has a long history of human impacts and environmental feedbacks. Historically, the human response to these feedbacks has generally not been in a timely fashion, but rather has been motivated by dire conditions and directly addressed only to those conditions. The Dust Bowl, for example, led to adoption of soil management practices to reduce erosion throughout the Midwest, but little was done to address underlying water quality issues, although the development of the hypoxic "Dead Zone" in the Gulf of Mexico has brought water quality issues to the forefront. The new biofuels mandates prompting another land use change in this region have provided an opportunity to examine the problem from a broader perspective, considering the trade-offs and possible solutions before a new set of dire conditions develops. Because the history of these two watersheds has highlighted the importance of environmental feedbacks and the need for a timely human response, we propose to use a watershed model in an integrated systems model framework that allows for feedbacks between the human and the environmental systems. In this way, human actions impact the environment; these impacts in turn feed back onto the human system, which then must respond or find a solution. The observed data, then, is used not only for calibration, which can give the "right" answer for the wrong reasons (Kirchner, 2006), but also to gain a more fundamental understanding of the underlying processes controlling the catchment response (Klemeš, 1988).

3.6 Conclusions

This paper has presented a comparative study between two adjacent watersheds in the Midwestern US. Despite their close geographical proximity, the similar size of their drainage areas, and experiencing in general the same temperate climate, significant hydrologic differences, were found both between the watersheds and within them. These differences are the result of both natural and human impacts to the region, with the major geologic difference being glacial history and the major anthropogenic differences being reservoirs, tile

drainage, and land development. Although the general climate is similar, spatial variation in precipitation further increases the heterogeneity of the hydrologic response. The differences highlighted here are the result of the co-evolution over time of the human and natural system and are thus relevant to any future changes that may be imposed on this coupled system.

Analyses of hydrologic signatures have revealed three main controls on the hydrologic response of these two watersheds. First, the precipitation inputs themselves are spatially and temporally variable in this region, and this area effect can be seen in both the annual and average monthly precipitation. At smaller time scales, this spatio-temporal variability can increase the heterogeneity of the catchment response due to the intensity of small convective storms common to this region. Second, the storage effects of inline reservoirs play a role in the catchment response of both watersheds. The storage in two large, main-stem, flood control reservoirs in Kaskaskia is instrumental in reducing NO_3-N output from the watershed as well as reducing flow variability on controlled reaches, while the storage in smaller, municipal water supply reservoirs in Sangamon is diverted for human use, further reducing low flows and increasing the number of zero-flow days expected for the size of the drainage area. Third, at the scale of this study, tile drainage plays an extensive role in the catchment response in both watersheds, especially in the Sangamon, where a higher proportion of the land is tile drained. These effects manifest as a homogeneous regime curve, increased heterogeneity in the low flows, and a higher BI relative to less-tiled areas. However, because tile drainage in this region is also generally associated with intensive row-crop agriculture, it may be the combined effect of these land modifications that is being observed.

Analysis of the histories of the case study catchments shows that modifications and impacts have been layered on top of each other through time: fire, prairie conversion, surface drainage, subsurface (tile) drainage, reservoirs, intensification of row cropping, erosion, fertilizer application, etc. Spatially, this layering does not always follow watershed boundaries, but often human ones; resulting in the formation of a different mosaic of layers of change in each watershed. These mosaics of change, combined with the controls identified by analysis of the hydrologic signatures, is manifested in the summer low flow behavior, where greater variability in duration, frequency, and relative magnitude is associated with streamflow deficits in Sangamon compared to those in Kaskaskia. These differences would all affect

the suitability of certain subcatchments for growing *Miscanthus*, and thus must be taken into account in future plans for biofuels expansion to avoid worsening or creating water stress conditions.

Acknowledgments

The research presented here was done as part of the NSF-funded project Interdependence, Resilience, and Sustainability of Infrastructures for Biofuel Development (NSF grant EFRI-083598, X. Cai, PI). We also wish to thank our anonymous referees for their insights and helpful criticism which have helped us improve this work.

Chapter 4

Water for food, energy, and the environment: projecting the trajectory of flow regime and water quality with increasing biofuel crop production in the US Corn Belt

Abstract

To better address the dynamic interactions between human and hydrologic systems, we develop an integrated modeling framework that employs a System of Systems optimization model to emulate human development decisions which are then incorporated into a watershed model to estimate the resulting hydrologic impacts. The two models are run interactively to simulate the co-evolution of coupled human-nature systems, such that reciprocal feedbacks between hydrologic processes and human decisions (i.e., human impacts on critical low flows and hydrologic impacts on human decisions on land and water use) can be assessed. The framework is applied to a Midwestern US agricultural watershed, in the context of proposed biofuels development. This operation is illustrated by projecting two possible future co-evolution trajectories, both of which use dedicated biofuel crops to reduce annual watershed nitrate export while meeting ethanol production targets. Imposition of a primary external driver (biofuel mandate) combined with different secondary drivers (water quality targets) results in highly nonlinear and multi-scale responses of both the human and hydrologic systems, including multiple trade-offs, impacting the future co-evolution of the system in complex, heterogeneous ways. The strength of the hydrologic response is sensitive to the magnitude of the secondary driver; 45% nitrate reduction target leads to noticeable impacts at the outlet, while a 30% reduction leads to dominant impacts that are mainly local.

This work has been submitted for publication to Water Resources Research as: Yaeger, M A, M Housh, X Cai, and M Sivapalan, Water for food, energy, and the environment: projecting the trajectory of flow regime and water quality with increasing biofuel crop production in the US Corn Belt. All figures and tables were created by Mary Yaeger unless otherwise indicated.

The local responses are conditioned by previous human hydrologic modifications and their spatial relationship to the new biofuel development, highlighting the importance of past co-evolutionary history in predicting future trajectories of change.

4.1 Introduction

Humans and hydrology are strongly interconnected (Falkenmark, 1997). With growing populations and improving technologies, the human impact on hydrology has left very little of the environment untouched (DeFries and Eshleman, 2004; Vörösmarty et al., 2010; Carpenter et al., 2011). Thus it is more important than ever to find some sustainable balance between the water needs of humans for food and energy and those of the environment (Cai et al., 2002, 2003; Baron et al., 2002; Falkenmark, 2003; DeFries et al., 2004; Rockström et al., 2009; Perrone and Hornberger, 2014). There are serious challenges to this goal, notably the complexity of the interactions between the human and hydrologic systems as they undergo continuing or accelerating changes due to both human and natural external drivers (Thomas, 2001; Parker et al., 2008; Wagener et al., 2010). As the two systems co-evolve through time, this propagation of external driving forces through the system is manifest in heterogeneous ways, with the complex, nonlinear response leading to emergent dynamics (Phillips, 2001; Liu et al., 2007a). While current research has provided insights into either the impacts humans have on hydrology or, conversely, hydrological variability on human society, the dynamic nature of these interactions are still not well understood (Collins et al., 2010; Wagener et al., 2010). Sustainable management of water resources in an uncertain future requires long-term prediction of alternative states of these complex systems. Traditional methods, such as the employment of multiple perfect foresight scenarios, may be inadequate for this task and thus new approaches are needed (Wagener et al., 2010; Thompson et al., 2013).

The difficulty presented by provides the motivation for the current research, which develops a quantitative framework that holistically considers the human and hydrologic systems as a linked, co-evolving system. Because agriculture, among all human factors affecting hydrology, is a major global vehicle for land use change (Gordon et al., 2008), accounting

for about 85% of world consumptive water usage (Foley et al., 2005), and is thus a primary driving force of these systems, the framework was applied to an agricultural system which has undergone extensive land use change in the past and for which new biofuel development has been proposed. Since the historical trajectories of these coupled systems can provide a better understanding of the legacy effects of past changes (Yaeger et al., 2013) and possible emergent behavior due to human modifications (Wagener et al., 2010), the historical co-evolution of the case study region is presented next as a specific example of the general problem motivating the work.

This co-evolution is very clearly evident in the history of the central Midwestern US, which has been transformed over time from a wet perennial grassland system to an artificially-drained, intensively cultivated agricultural system. The conversion of the natural vegetation cover into annual crops and associated drainage activities mobilized both sediments and nutrients, leading to significant soil erosion and loss of soil fertility (David et al., 2001; Turner and Rabalais, 2003). In response to these feedbacks, humans improved soil management and reduced soil losses; however, the development of inexpensive nitrogen fertilizers allowed them to compensate for soil nutrient losses (Turner and Rabalais, 2003). The continued expansion of subsurface drainage, intensification of agriculture, and increasing fertilizer application worsened the nutrient leaching, leading eventually to water quality problems, both at local (Smith et al., 1990) and very large scales (Goolsby et al., 1999). The human system response to these hydrologic feedbacks at the local scale has been adaptation using water treatment technology and at the large-scale, mitigation using best management practices (BMPs) and policy recommendations (David et al., 2013). The result at the large scale has not been as good as expected (Sprague et al., 2011) due in part to the spatial disconnect between the source of the impacts (Midwest) and where the feedbacks manifest (Gulf of Mexico) (David et al., 2013), compounded by unintended consequences of bioenergy policies (Donner and Kucharik, 2008).

The future status of the environment, especially the hydrological aspects, is now challenged by energy development plans in the U.S. Midwest. In 2007 US bioenergy policies were updated, with the Renewable Fuel Standard (RFS2) expanded to require 36 billion gallons of renewable fuels be produced by 2022, more than half of which were to come from cellulosic

sources. Current biofuel production consists mainly of corn-based ethanol; the increased demand from these mandates led to greater corn production and thus increased fertilizer application throughout the Midwest, which resulted in the unintended conflict with nutrient reduction efforts (David et al., 2013). Cellulose-based ethanol is not yet in large scale production; however, high-yielding perennial grasses such as *Miscanthus* and switch-grass are under serious consideration as biofuel feedstock crops (Jain et al., 2010; Martin, 2011). For simplicity, because in the study area *Miscanthus* has a much higher per hectare yield compared to switch-grass, this paper will henceforth focus on *Miscanthus* and its effects. From the perspective of linked, co-evolving human-hydrologic systems, the changes to be imposed by the RFS2 mandate represent a new external driver, the impacts from which may manifest as possible solutions to existing problems as well as potentially creating new problems.

The dual nature of the potential impacts of meeting the cellulosic mandate with *Miscanthus* emphasizes the importance of considering the system as a linked whole, given the unintended outcomes already experienced. On the one hand, *Miscanthus* could provide solutions to both the nitrate export problem (Ng et al., 2010; Vanloocke et al., 2010) and the food vs. fuel debate (Dohleman et al., 2010). On the other hand, there are trade-offs: while *Miscanthus* has lower fertilizer requirements than corn, its longer growing season results in higher water usage (McIsaac et al., 2010; Vanloocke et al., 2010; Le et al., 2011), reducing soil moisture storage and hillslope drainage, and in this region, summer low flows as well. The potential for this reduced streamflow to create new problems increases if other human adaptations such as water supply reservoirs, etc., are affected. Furthermore, the qualitative analysis of this regions historical trajectory has shown that cross-scale impacts and feedbacks between the human and hydrologic systems are important here. The large-scale feedback is from the natural system far downstream in the Gulf of Mexico, but it is the human system upstream in the Midwest that needs to take action (David et al., 2013). This would require institutional intervention, since this feedback is not felt locally and thus the human system has insufficient motivation to change its behavior.

The example of the general problem just presented illustrates that while we may be able to identify and qualitatively understand bi-directional feedbacks in a specific case, a key issue

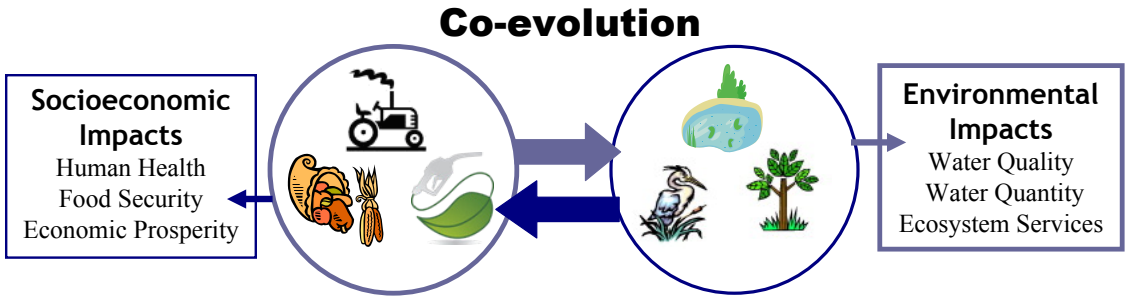

Figure 4.1: Conceptual model of co-evolution and competition between human and environmental systems

is how to quantify them, and in a general way. Concepts from recent work provide a basis for accomplishing this. The first considers watersheds as coupled human and natural systems (CHANS) which should be treated as a whole in order to gain deeper understanding of how the two systems coexist and interact (Liu et al., 2007b). This idea is illustrated conceptually in Figure 4.1, where the human and natural systems are shown in terms of their competing water needs, with water for food and energy on the left, and water for the environment on the right. The large arrows represent the interactions between the two systems; the small arrows represent the impacts of each system on the other. These interactions arise from internal responses to external drivers propagating heterogeneously through the coupled system and which, over time, manifest in complex, nonlinear ways, e.g. time lags, local and cross-scale impacts, resilience, and legacy effects (Liu et al., 2007a). Thus watersheds are CHANS co-evolving through time and should be treated in an integrated, dynamic way in order to identify the underlying fundamental principles of the coupled system (Sivapalan et al., 2012). This more open approach to traditional foresight scenario modeling allows for human impacts on the hydrologic system to influence or constrain further human decisions, enabling course corrections along the co-evolutionary trajectory, thereby providing predictive insights into the dynamics of the coupled system. In addition, these trajectories of change provide a way to understand and visually represent a constrained range of possible futures which can then be used to guide further analysis (Thompson et al., 2013). Here we may

highlight the following need: although CHANS is conceptually attractive for the problem specified above, a dynamic modeling framework that can facilitate the analysis is still needed for empirical findings in terms of the feedbacks and dynamics, which are not well presented in the literature.

To this end, we combine these concepts in the integrated modeling framework developed in this paper, wherein a decision-making model specifies the most economically beneficial spatial pattern of land use change for biofuels development, which is then incorporated in a watershed model to determine the hydrologic impacts that must be addressed by the human system. In this way, possible trajectories of system co-evolution visualize how externally-driven changes propagate through the system. Although the framework is general enough that it can be applied to many kinds of CHANS, it is demonstrated here by applying it to an agricultural watershed within the case study region described previously. The objectives of this paper are to develop an integrated modeling framework and apply it to the case study CHANS by projecting the co-evolution of the coupled system into the future in stages in order to understand how external drivers propagate through the coupled system, and in doing so identify some of the multiple trade-offs necessary at the watershed scale for sustainable water resource management in a biofuels development context.

This paper is outlined as follows: the first section has provided background for the problem and motivation for the particular approach. The second section conceptually describes the proposed interactive framework, while the third describes the pertinent characteristics of the case study area. Section 4.4 presents the application of the framework, including descriptions of the two models used. Section 4.5 presents the results of trajectories based on different human system goals, followed by a discussion in Section 4.6 of their implications with respect to the various trade-offs related the problem outlined in Section 4.1. The last section provides a summary of the main findings of the work and major conclusions.

4.2 Integrated modeling framework

Recent studies of agriculture-biofuels systems have become more integrated to better explore the trade-offs between water quality policies, environmental needs, and biofuels devel-

opment. Integrated hydro-agro-economic optimization models have been developed (Moraes et al., 2009) as an alternative to assuming that a certain percentage of the watershed in a hydrologic model undergoes biofuels development, either as a random percentage (Ng et al., 2010; Wu et al., 2012; Wu and Liu, 2012) or guided by other studies (Demissie et al., 2012). Other integrated studies implement commonly-used spatially distributed biophysical and biochemical models to determine soil moisture and water quality, with a separate economic model providing the location of the land use change (Zhang and Schilling, 2006; Egbendewe-Mondzozo et al., 2011). However, in these latter cases, multi-objective optimization is used to reduce environmental impacts, depending on policies in effect. A hybrid integrated modeling framework, where an economic model driven by external inputs such as ethanol prices provides location and amount of biofuel development to a watershed model (Secchi et al., 2011) gives some insight into how outside forces drive local change, which could be used to show how these changes propagate through the stream network. However, in the scenarios presented, the impacts are generally one way (human system on environment), and the output is the predicted hydrologic response. In the integrated modeling framework developed in this paper, the impacts are in both directions, and the internal response of each system to the external drivers is more dynamic, since the purpose is to mimic co-evolution in stages, rather than predict the state of the system at an endpoint as a result of predetermined scenario.

The integrated modeling framework itself is composed of two models, a systems optimization model representing the various interacting subsystems involved in biofuels development, and a watershed model representing the hydrologic model (Figure 4.2). Further details of each model are provided in Section 4.4. External drivers to the coupled system are both natural (climate) and human (economics, government policy, societal values). The state of the human system is assessed at each stage by spatial patterns of land use and total system profits. The condition of the hydrologic system affected by these land use patterns can be assessed both at the watershed outlet by total (NO_3-N) export and the 7 day minimum discharge, as well as spatially within the watershed by flow duration curves (FDCs), the threshold summer low flow (Q_{85}), or drought deficit analysis as outlined in Yaeger et al. (2013).

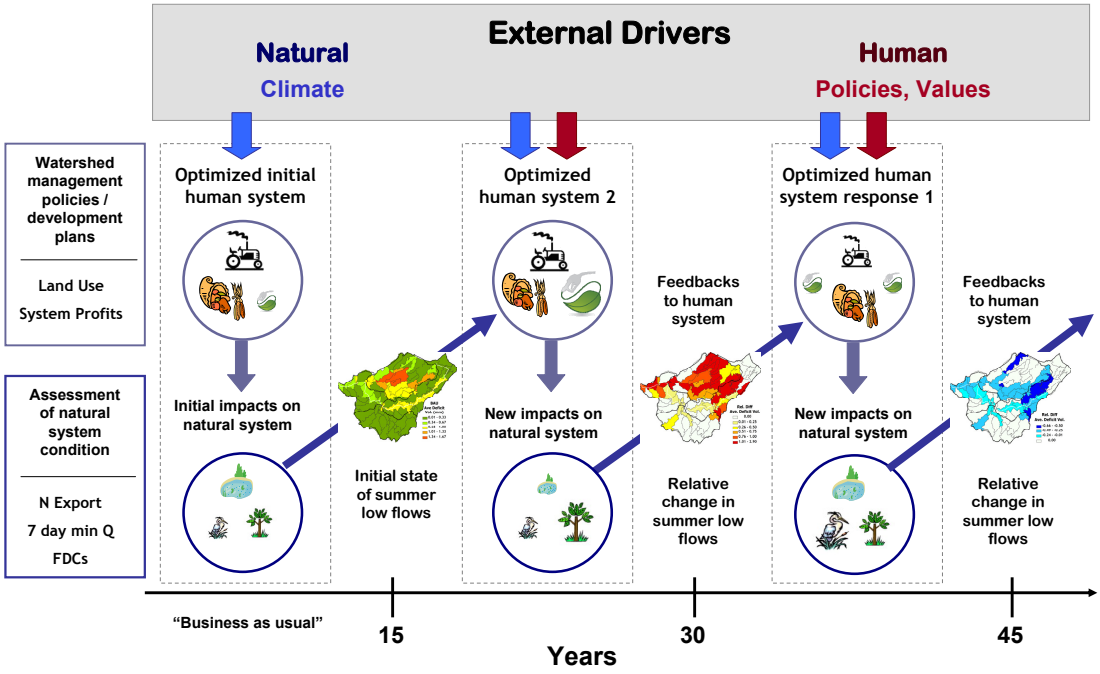

Figure 4.2: Conceptual model of the integrated modeling framework applied to show three stages of a trajectory of change. Note: for relative change inserts, red colors represent a decrease, blue colors an increase.

To create a trajectory of change, the two models representing the human and hydrologic systems are run interactively in tandem feedback control fashion, with the output from one model providing input decisions for the other. Here the optimal land use determined by the human system is implemented in the hydrologic model; resulting feedback from the hydrologic system is incorporated into the human system decision making process in the next stage of the trajectory. The state of each system is assessed at the decadal (10 – 20 year) time scale, which was selected because of the inertial nature of human system processes (Scheffer et al., 2003) as well as the time lags inherent in CHANS (Liu et al., 2007b), as some feedbacks take time to manifest while others may be precipitated by a change in climatic conditions such as an extended drought period. The arrows in Figure 4.2 represent the flow of impacts, where the lighter blue shows human impacts on hydrology, and the dark blue represents the feedback of hydrology onto the human system. Since this framework is only a first step towards the development of a fully coupled modeling system, the human responses

to the feedbacks from the hydrologic system are currently determined by the user of the framework. These responses are motivated in part by what outcomes society finds most desirable, or what problems it wishes to solve. The resulting decisions are then incorporated into the framework at the beginning of the next stage and can be altered over time, thus allowing for changes in direction as the possible consequences are explored.

To provide a baseline for comparison, each trajectory starts with optimized current conditions for the human system and the resulting impacts of these on the hydrologic system – the "Business as usual" box in Figure 4.2. Once the state of the coupled system under this representation of current conditions is established, a new external driver can be imposed and relative change in the assessment parameters can then be calculated. At the beginning of the second stage, a new external driver acts on the system, imposing changes such as bioenergy mandates or environmental policies that the human system must respond to, according to internal values and goals. Economic profits for the human system are then optimized for these new conditions and the resulting land use patterns are implemented in the watershed model to determine the impacts on the hydrology for the second decadal period (Figure 4.2, center box). In this way, potential feedbacks from the hydrologic system can be detected that might otherwise have gone unnoticed and timely action taken, if necessary, in Stage 3 (Figure 4.2, right box). Potential responses of the human system to these feedbacks depend both on the various trade-offs involved as well as on societal priorities. The conceptual trajectory ends here, but depending on the case study, further response may be necessary, in which case new decisions would be implemented in Stage 4, and so on. Multiple trajectories can be generated for each case study, depending on the possible responses to feedbacks arising from the initial changes imposed on the system.

4.3 Case Study

Hydrologic characteristics of the watershed

The 15,000 km^2 Sangamon River watershed is located in central IL, in the highly productive agricultural region known as the "Corn Belt" (Figure 4.3) and is a tributary of the Illinois

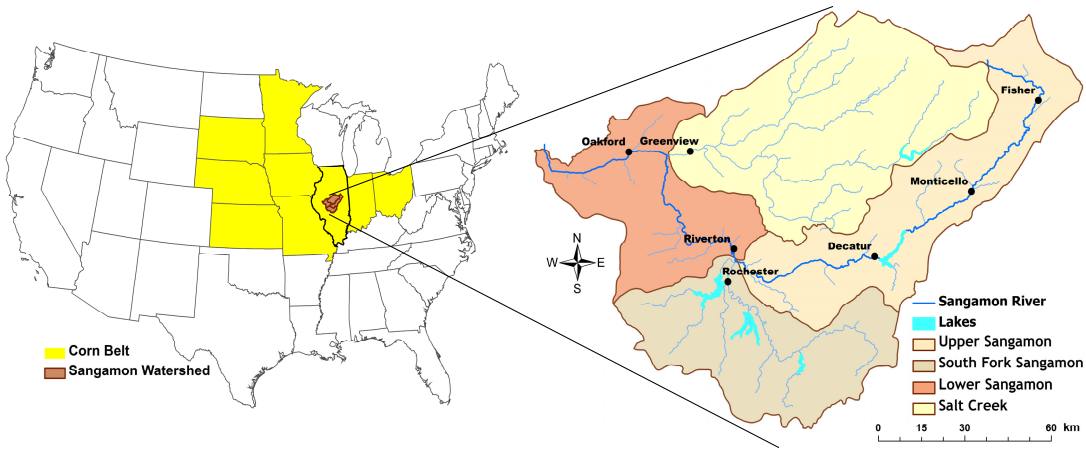

Figure 4.3: Map of the study site, showing major hydrologic units, lakes, and gauge locations for reference.

River which itself is a tributary of the Upper Mississippi River. Hydrologically, the Sangamon River is composed of an upper and lower main reach and two large tributaries Salt Creek in the north and South Fork Sangamon in the south (Figure 4.3, right panel). The topography is flat and covered by a deep layer of rich soil underlain by an impermeable hardpan layer. Within the Salt Creek and Upper Sangamon sub-watersheds is a topographical divide from a flat upland plateau in the northeast and a flat lowland region encompassing the South Fork and lower Sangamon. These geological features, combined with a humid climate would result in a poorly-drained landscape without an extensive artificial drainage system. The current streamflow regime is seasonal: the flood peak occurs in May and the lowest flows occur in late summer just after the peak of the growing season. During this time, smaller reaches often experience intermittent flow (Yaeger et al., 2013).

The climate in this region is humid and temperate, with an average of about 1000 mm of precipitation annually, although severe, extended droughts occur periodically (Woodhouse and Overpeck, 1998). Precipitation is not distributed uniformly but highly heterogeneously, both spatially and temporally. At the annual scale, this variation is seen both between years and between locations in the watershed (Figure 4.4). On average, the southern portions of the watershed (Springfield and Taylorville in Figure 4.4) receive more annual rainfall

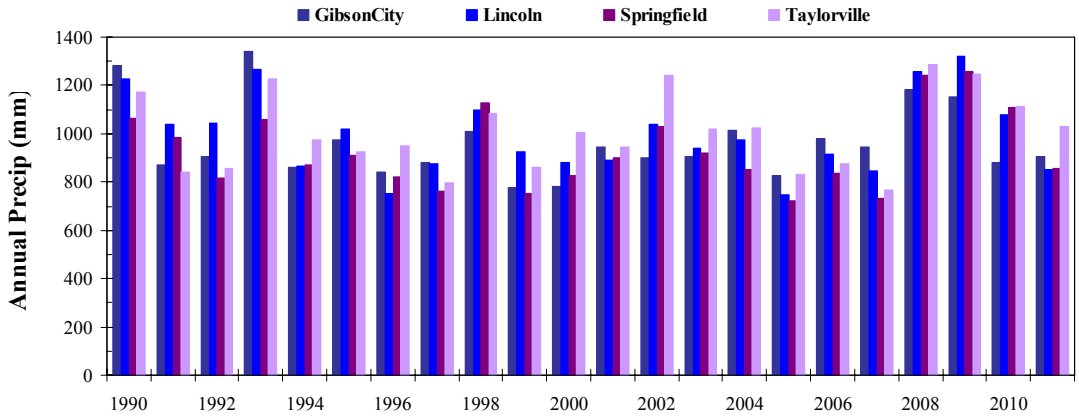

Figure 4.4: Annual precipitation at four locations within the watershed (from PRISM) showing interannual and spatial variability over the last 20 years.

than the north (Yaeger et al., 2013). Precipitation also varies seasonally within the year, with April and May generally being the wettest months and September and October the driest. Summer precipitation tends to be localized, from intense convective storms, and this spatial heterogeneity at small scales interacting with and filtering through the historical modifications to the system, increases the complexity of the catchment hydrologic response.

4.3.1 Human influences on hydrology

Because of the humid climate and flat watershed topography underlain by a hardpan subsoil layer, artificial drainage is necessary to maintain an unsaturated root zone during the growing season. About 90% of the watershed area is agricultural, consisting mainly of the annual row crops corn and soybeans, with the remaining 10% comprising a few small urban areas and some riparian forest. To support these intensive farming practices, the subsurface has been extensively tile-drained, with the surface drainage enhanced by ditching and channelization of existing streams (Yaeger et al., 2013). These significant human modifications have changed the hydrologic response of the watershed; both the quality and quantity of streamflow have been affected. Surface runoff is now infrequent, occurring mainly in winter due to frozen soils and in spring due to saturation excess. Hillslope drainage of soil storage is the main source of baseflow, with the majority contributed by tile drainage (Li et al., 2010a). Tile

drains and surface ditching together reduce the residence time of water in the catchment, thus bypassing many of the natural biogeochemical processes that remove NO_3-N (McIsaac and Hu, 2004; Royer et al., 2006). Additionally, the largest nitrate mass export occurs during the largest flows (Borah et al., 2003), often coinciding with the spring fertilizer application (Royer et al., 2006). Together, the hydrologic variability and human modifications to the landscape combine to produce the large annual export of NO_3-N from this watershed. Thus, large scale planting of *Miscanthus* in Sangamon could provide a much-needed water quality benefit without reducing the agricultural socioeconomic benefit.

There are no major flood control dams in this watershed, but there are several small, uncontrolled, inline reservoirs that serve as municipal water supplies or cooling ponds for power plants. In the Upper Sangamon, Lake Decatur supplies water for the city of Decatur (Figure 4.3); due to its location in the upper reaches of the Sangamon River, which experience more variable streamflow than the lower reaches, it is more vulnerable to extended drought periods (Yaeger et al., 2013). Two other water supply reservoirs are Lake Taylorville near the South Fork headwater region, which provides water supply for the city of Taylorville, and Lake Springfield near the confluence of the South Fork with the lower Sangamon, which serves as the surface water supply for Springfield, IL. There are also two major cooling ponds in the Sangamon watershed: Clinton Lake in the upper reaches of the Salt Creek tributary and Lake Sangchris in the South Fork tributary. There are also several point source locations in the watershed with discharge permits, which are based on historical low flows. These locations, along with the reservoir drainage areas, are regions that are possibly vulnerable to flow reduction caused by large-scale planting of *Miscanthus*.

4.4 Application of the Integrated Modeling Framework

Within the case study CHANS, the focus of the human system is on agriculture and biofuels development, while in the natural system the focus is mainly on the hydrology, considering both water quality and water quantity.

4.4.1 Human system: system of systems model

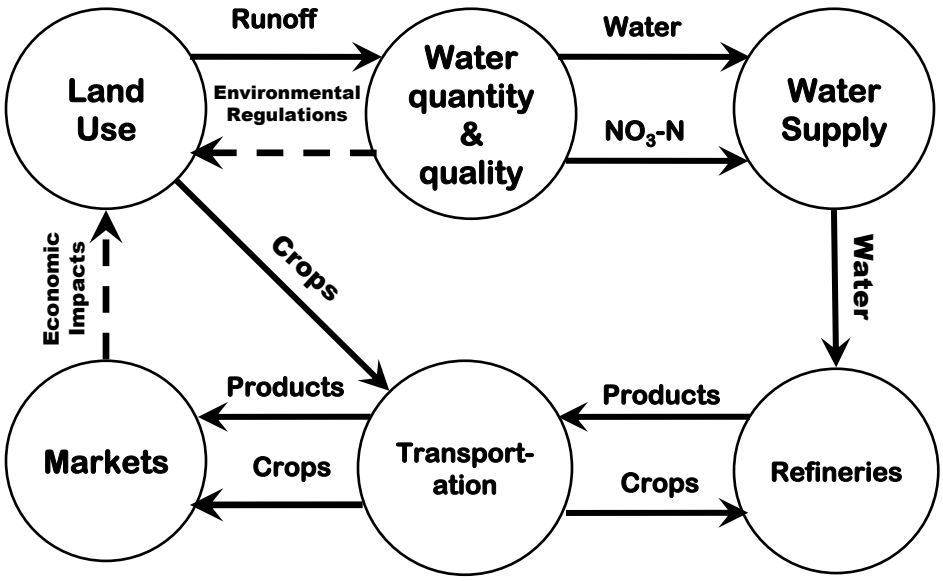

Figure 4.5: Schematic diagram of the SoS model, showing subsystem interdependencies (adapted from Housh, et al.(2014)).

The System of Systems model (SoS-Biofuel) (Housh et al., 2013) is used to represent the human system in the integrated framework. The SoS approach was chosen because it encompasses all subsystems in the biofuel development system (Figure 4.5), and includes transportation infrastructure, refinery infrastructure, and economic decisions as well as the land allocation decisions already discussed. In addition, the various subsystems are mutually interdependent, so that changes to or decisions made in one subsystem impact all other subsystems. The SoS model also includes a watershed module to encompass the relationships between environmental policy, water quantity and quality, and the various subsystems involved in biofuel development. The overall objective of the system is to maximize the net economic profit, considering revenues and costs (capital and operational) in each subsystem (Figure 4.5). This profit maximization is done under physical, operational and environmental constraints. The SoS approach, therefore, is able to simultaneously take into consideration land use allocation to various crops, infrastructure needs of transportation and refineries, as

well as environmental requirements.

The SoS-Biofuels model was developed for the case study watershed using data and modeled inputs from groups representing the various subsystems, including hydrology, economics, transportation, refinery engineering, and agriculture (Housh et al., 2013). The version of the model used in this study uses a 12 year rolling horizon to determine the conditions for the average optimal year and has previously been applied to the Sangamon watershed to explore the impact of biofuels land use change on the hydrologic system as well as the impact of water quality targets on the biofuel production system as a whole (Housh et al., 2014).

4.4.2 Natural system: watershed hydrologic model

While the SoS model includes a watershed module to reflect the environmental targets within the development, it is a simplified representation of the watershed and is not intended for performing detailed spatial and temporal analysis of the hydrological responses of the watershed. Therefore, within the integrated modeling framework, the Soil and Water Assessment Tool (SWAT) version 2005 developed by the US Department of Agriculture Agricultural Research Service (USDA ARS) was used to provide detailed responses of the hydrologic system resulting from decisions of the human system (represented by SoS) (Arnold and Fohrer, 2005; Gassman et al., 2007). This model was chosen because it is well-documented, required input data is readily available, and it has been used extensively in similar biofuels development applications in the Upper Mississippi River Basin (Secchi et al., 2011; Demissie et al., 2012; Wu et al., 2012; Wu and Liu, 2012), in the Raccoon River watershed in Iowa (Schilling et al., 2008), and the Salt Creek tributary of the Sangamon (Ng et al., 2010). Of these, Ng et al. (2010) developed SWAT crop growth parameters for *Miscanthus* which we have used in the simulations presented in this paper.

The SWAT model of the Sangamon River watershed was developed using the input data presented in Table 4.1. Current local agricultural practices were modeled after the land management scheme presented in Hu et al. (2007), while ranges for other parameter values were obtained from the literature cited above, especially Ng et al. (2010), since the Salt Creek is part of the Sangamon. In total, the watershed was delineated into 104 sub-watersheds,

Table 4.1: Description of and sources for data used in this study

Data Description		**Data Source**
30-m Land cover, 2001 30-m digital elevation map (DEM)	USGS[1] National Map	http://nationalmap.gov/viewer.html
STATSGO soils map, by state	USDA NRCS[2]	http://soildatamart.nrcs.usda.gov/USDGSM.aspx
Daily precipitation	NOAA[3] NCDC	http://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn
Daily streamflow (calibration)	USGS NWIS	http://maps.waterdata.usgs.gov/mapper/index.html
Annual Crop Yields, by state and county (calibration)	USDA NASS	http://quickstats.nass.usda.gov/

[1]US Geological Survey, [2]Natural Resources Conservation Service, [3]National Oceanic and Atmospheric Administration

ranging in area from 20 km^2 to nearly 100 km^2. Due to the homogeneity of land use and management practices in this watershed, and following Ng et al. (2010), multiple HRUs per sub-watershed were not implemented.

The model was run for a total of 12 years (1992 to 2003) chosen because they incorporated a range of dry and wet years. The calibration and validation of the finished model was carried out according to the procedure outlined in Neitsch et al. (2010) for daily streamflow at four points along the main-stem and one at the outlet of each of the two main tributaries, and monthly NO_3-N load at the outlet to Salt Creek. The Nash Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), root mean squared error standard deviation ratio (RSR), and percent bias (PBIAS) were used to determine goodness of fit between modeled and observed data as outlined in Moriasi et al. (2007). For the calibration period (1998 to 2003), the NSE for monthly discharge, ranging from 0.70 – 0.90, is considered "very good" (Moriasi et al., 2007) while those for daily discharge (0.55 – 0.73) have been reported as "good" in other papers [e.g. Van Loon and Van Lanen (2013)]. In addition, RSR (0.31 – 0.55) and PBIAS (-14.66 – 1.89) for monthly discharge all fell within the "good" to "very good" category according to Moriasi et al. (2007), and the values of both metrics for daily discharge a fell into very similar ranges. For the validation period (1992 to 1997) the NSE for monthly discharge (0.62 – 0.84) was slightly lower, but still within the "good" to "very good" classification, while

for daily discharge (0.33 – 0.61) it was slightly lower than the median but still within values in published literature (Moriasi et al., 2007). In general, model performance improved with increasing drainage area, with the exception of the tributaries; here the model captured the flow from South Fork much better than that from the Salt Creek, which is larger.

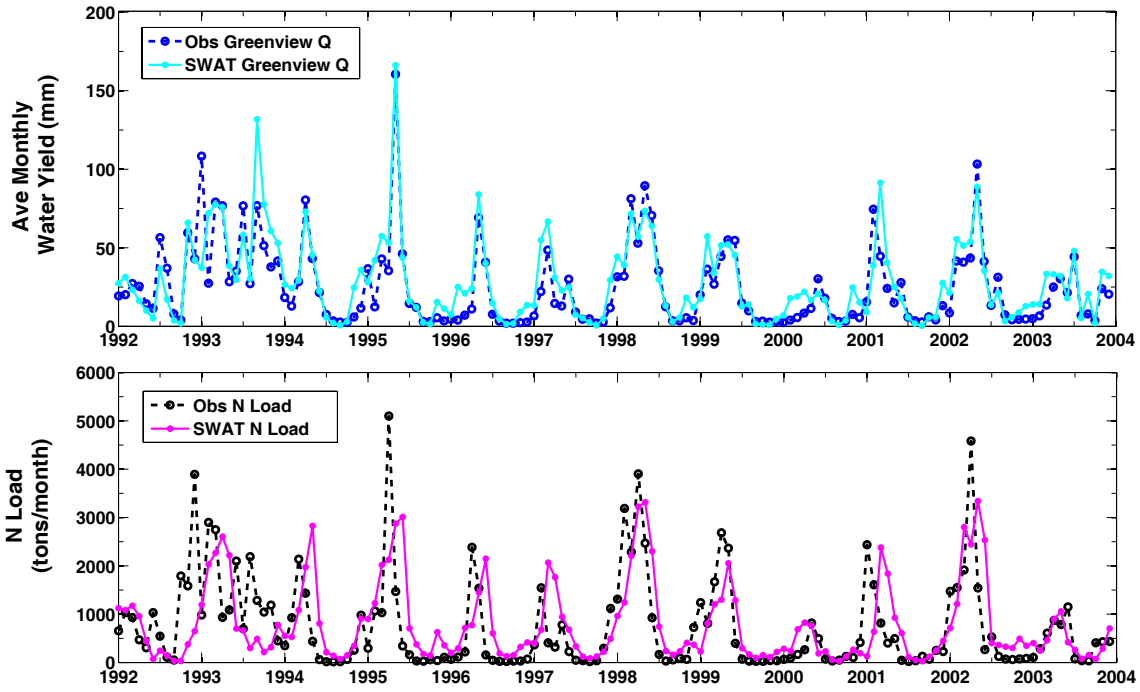

Figure 4.6: Time series of simulated and observed monthly water yield and NO_3-N load at the water quality calibration point at the outlet to Salt Creek.

Although the calibration period NSE for monthly NO_3-N was low (0.40), it was well within the range published in the literature (Moriasi et al., 2007); additionally, PBIAS was -7.64%, which they consider very good for a monthly time-step. For the validation period, PBIAS was still very good, but NSE was low (-0.09), likely due to missed nutrient pulses during the first few years of the simulation (Figure 4.6). While the general seasonal patterns of streamflow and NO_3-N load were captured by the model, it tends to underestimate the largest load peaks. Interestingly, at the Greenview location, the model does capture the largest streamflow peak (Figure 4.6, top panel), but in general, it underestimates very large flows at the other locations, leading, in part, to underestimation of the largest loads.

Because low flows were of interest in this particular case study, the model was calibrated

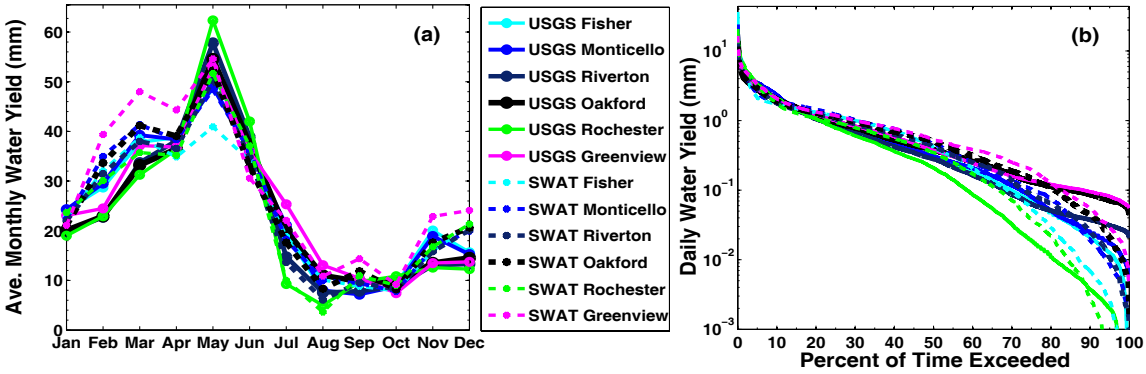

Figure 4.7: Simulated (dashed lines) and observed (solid lines) monthly regime curve (a) and daily flow duration curve (b) for 6 calibration/validation locations in the watershed.

with care to capturing the daily flow regime. In addition to fitting the daily hydrograph, the model output was also checked against the monthly regime curve (RC) and daily flow duration curve (FDC) to ensure that the model could adequately represent the low flow patterns needed for the low flow analysis (Figure 4.7a and 4.7b, respectively). These figures show the model tends to overestimate winter flows, and underestimate the spring flood peak, but overall, the general flow regime patterns are captured. The goodness-of-fit metrics showed consistent poor performance at the Greenview location, which is just upstream of the Oakford gauge. This may be partly due to the unusually high low flows from this watershed, resulting from treated wastewater discharge into Sugar Creek, a tributary of the Salt Creek in the northern part of the watershed, from sources outside the Sangamon (Yaeger et al., 2013) that are not currently in the model.

4.4.3 Coupled system: interactive modeling within the framework

The decadal analysis interval presented conceptually in Section 4.2 is applicable to this case study because the lifespan of a *Miscanthus* field is estimated to be about 15 – 20 years. This allows for a more quick response should environmental problems arise, and also takes into account farmers' unwillingness to abandon a long term crop for which they have had to wait 3 years to become fully established. For this demonstration of the framework, a

12 year analysis interval has been adopted to better match the 12 year rolling horizon in the optimization model, but the framework itself is flexible enough to accommodate longer periods.

The optimized solution that the SoS model provides is the state of the human system for a typical year in the 12 year horizon. This solution has taken into account averaged data within those 12 years; thus when the land use solution is incorporated into the hydrologic model it is applied uniformly over the entire analysis interval. Moreover, in this case study the pattern of wet and dry years seen in Figure 4.4 is assumed to repeat into the future for the trajectories presented and discussed here. However, the framework can easily accommodate synthetic precipitation time series where precipitation is increasing or decreasing with time, or where one stage is dominated by a prolonged drought period.

The spatial scale of the SoS model is a 100 km x 100 km grid, and the temporal scale is monthly; the spatial scale of the SWAT model is the sub-watershed, which can vary in size up to 100 km^2, while the temporal scale is daily. Both models incorporate the same land uses and management practices: a corn/soy row crop rotation, fallow/CRP conservation land, urban area, and *Miscanthus*. In the SoS model, multiple land uses are modeled within each land parcel as a proportion of the total parcel area, while within the SWAT model, the predominant land use in each sub-watershed is assumed to cover the entire sub-watershed area. Thus, in the SWAT model, *Miscanthus* only appears in sub-watersheds where it is the predominant land use. This preserves the general spatial clustering of *Miscanthus* within the watershed, and thus its proximity to refineries and transportation networks which are an essential part of the SoS model.

4.5 Results

The trajectories presented in the following section are based on the premise that humans highly value reduction of NO_3-N export to the Gulf of Mexico; this then becomes a secondary external driver imposed on the coupled system, directly affecting both the human system response to the primary external driver (the RFS2 mandate) as well as the resulting feedbacks from the hydrologic system. The initial state of the coupled system in both cases is an

optimized land use solution provided by the SoS model for the current cropping system and corn ethanol production, hereafter referred to as "business as usual" (BAU); the state of the hydrologic system at the watershed outlet after a 12 year period of these conditions provides a baseline to which future impacts can be compared. At the end of this time, the RFS2 cellulosic ethanol mandate is imposed along with an initial NO_3-N reduction target, so that subsequent stages of the trajectories form consecutive impact/feedback loops of internal responses to these external drivers.

4.5.1 Trajectory 1: Reduce annual watershed NO_3-N export by 45%

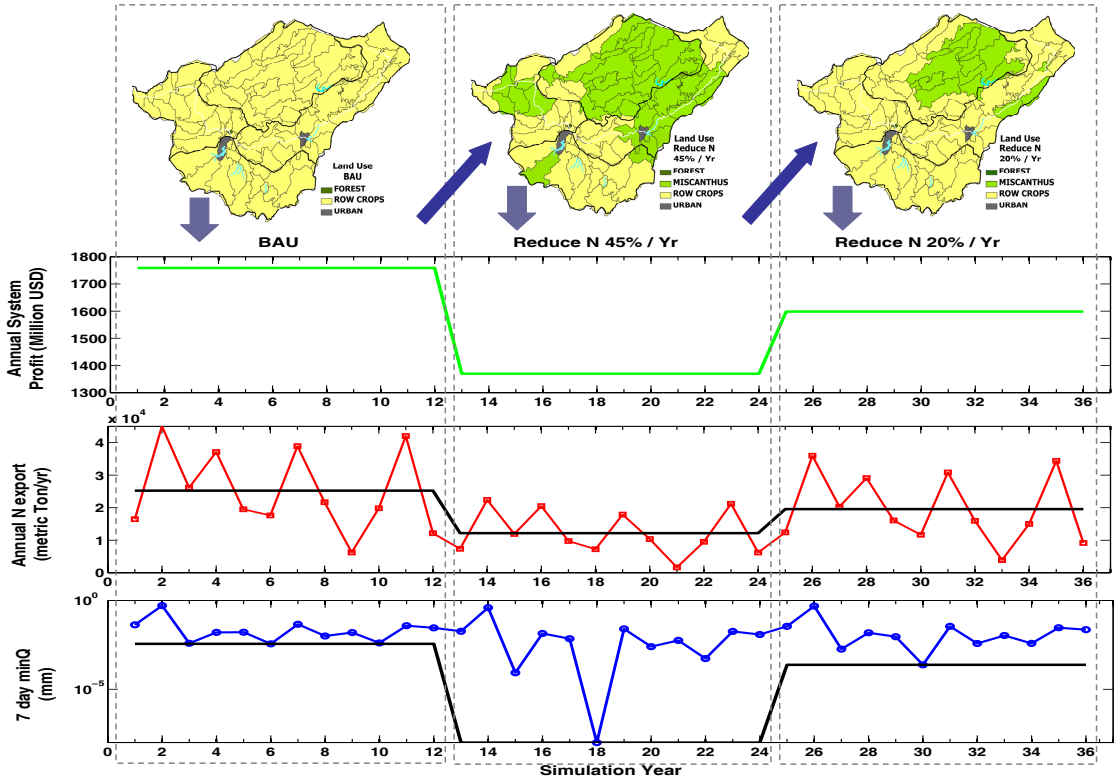

Figure 4.8: Trajectory 1, where the initial goal is to meet EPA recommended 45% reduction in annual N export from the watershed. Note: land use maps, annual system profit, annual watershed N export, and annual 7-day minQ at outlet, are shown top to bottom, respectively; straight lines represent mean annual values for profit and N, and the minimum value for 7-day minQ.

The initial condition of the system is shown in Figure 4.8 (left box). Land use is row crop corn/soy rotation, and the resulting NO_3-N yearly export from the watershed is fairly high

(about 25,000 metric tons/ yr on average). Within this period there are three very dry years, as shown by the 7 day minimum flows, where the annual minima are close to the minimum for the entire period. Because this region has been highlighted as the source of the majority of NO_3-N export to the Gulf of Mexico, one possible value for the secondary driver imposed in Stage 2 is 45%, recommended by the EPA for total NO_3-N export to the Gulf, but which could also be applied at the watershed level (Figure 4.8, center box). This requires about half of the agricultural land area to be converted to *Miscanthus*, with the majority of this conversion occurring in the high-yield tributaries Salt Creek and Upper Sangamon (in the drainage area of Lake Decatur), as well as small regions in the Lower Sangamon and in the drainage area of Lake Springfield (Figure 4.8). In addition, the average annual system profit decreases by 22% from the BAU case. The impacts on hydrology are mixed; while the water quality benefits are large, with a 47% reduction in watershed NO_3-N export annually, the 7 day minimum flow at the outlet for the second 12 year period is nearly 0, with lower annual 7 day minima during dry years than under the BAU condition. Within the watershed, inflow to two of the municipal water supply lakes and one of the cooling reservoirs is reduced due to extensive *Miscanthus* planting upstream.

Given the negative impacts on the human system (loss of agricultural land and reduction in profits) and the mixed feedbacks from the hydrologic system (reduction in both NO_3-N export and low flows at the outlet and within the watershed), there are many possible responses the human system could have and multiple trade-offs to consider. Nearly half of the agricultural land was converted to dedicated biofuels crops, which might be a concern for food security if more watersheds followed suit; this land conversion also resulted in widespread flow reduction that was evident even at the watershed outlet. Thus one plausible response might be to relax the NO_3-N reduction requirement in an attempt to find some balance between competing land uses (food or fuel) and competing water needs (quality and quantity), as illustrated in Figure 4.8 (right box), where the NO_3-N export reduction requirement is now only 20% of the BAU export.

Achieving this new water quality goal allows some of the land previously converted to *Miscanthus* to revert back to row crop agriculture, leaving 20% of the total agricultural land area under *Miscanthus* cultivation and resulting in a 17% increase in average annual

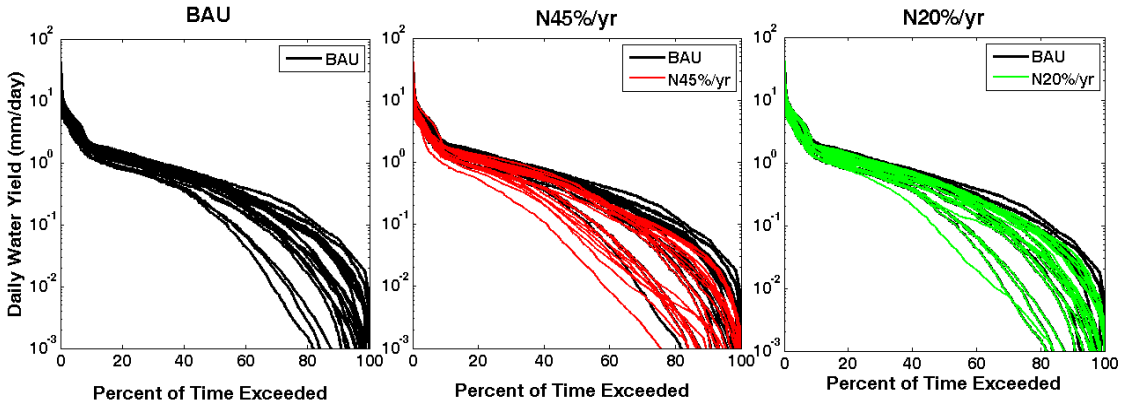

Figure 4.9: Flow duration curves for selected subcatchments within the watershed for the three stages of Trajectory 1. Subcatchments shown are up- and downstream of lakes, at gauge points on the mainstem, and at the outlet.

profit. Because the costs associated with production of cellulosic ethanol from *Miscanthus* are greater than those for corn, the total profits are still less than those from the BAU condition. The increase in row-crop cultivation also results in higher NO_3-N export from the watershed, although it is still less than under the BAU land use. At the same time, the 7 day minimum flow for the last 12 year period increases nearly to what it was prior to the initial conversion to *Miscanthus*. This small impact on NO_3-N export but large impact on flow just described illustrates some of the nonlinearity in the hydrologic response to the external changes. Within the watershed, due to the various interdependencies considered by the SoS model, the optimal locations for the remaining *Miscanthus* are now concentrated in a smaller area in Salt Creek and a small portion of the Lake Decatur drainage area in Upper Sangamon. The net result of this is a nearly complete restoration to historical inflows to both water supply lakes, but only partial restoration of historical inflows to the large cooling pond in Salt Creek. This effect is illustrated by selected FDCs for the entire trajectory (Figure 4.9). These sub-watersheds represent the six locations shown in Figure 4.3, as well those upstream and downstream of the lakes shown in Figure 4.3 and described in Section 4.3.1. The sub-watersheds most severely affected by the land use changes show a distinct shift to the left compared to their state under the BAU condition (Figure 4.9, center panel); not only are the low flow tails affected, but the average flows of the middle limb are also reduced. When the percentage of *Miscanthus* in the entire watershed is reduced in response

to this feedback, there is a corresponding shift back to the initial condition (Figure 4.9, right panel); however, some sub-watersheds are still affected, and their flow remains reduced.

4.5.2 Trajectory 2: Reduce annual watershed NO_3-N export by 30%

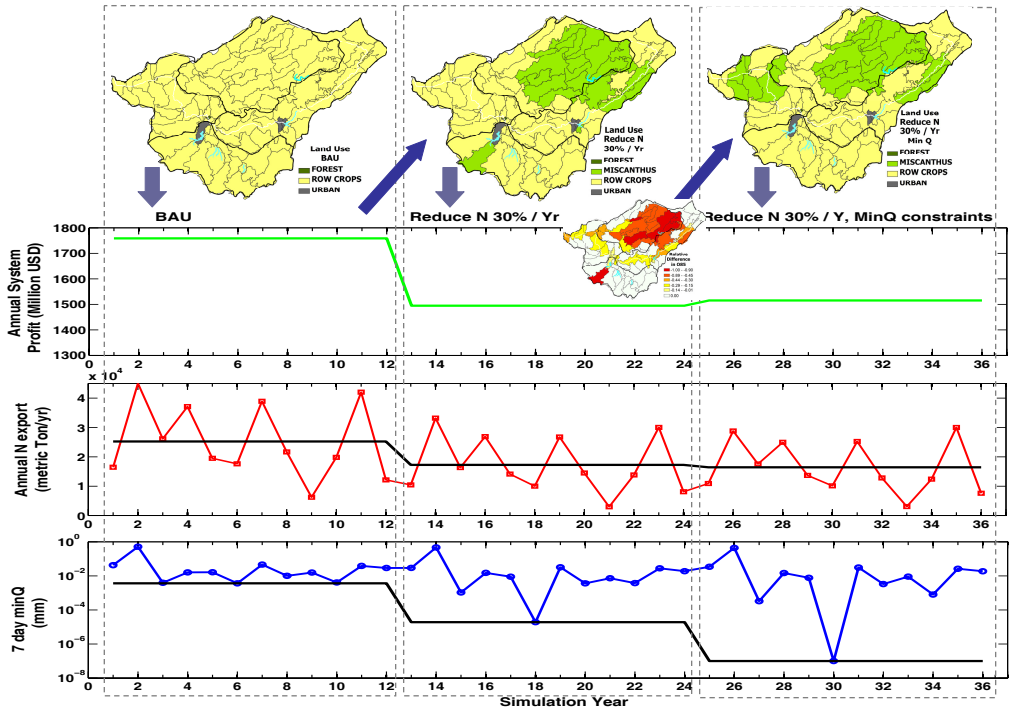

Figure 4.10: Trajectory 2, where the initial goal is to reduce watershed NO_3-N export by 30%; the insert shows the hydrologic feedback given as relative change in Q_{85} flow (from the flow duration curve). Note: land use maps, annual system profit, annual watershed N export, and annual 7-day minQ at outlet, are shown top to bottom, respectively; straight lines represent mean annual values for profit and N, and the minimum value for 7-day minQ.

In the second system trajectory, the initial conditions and the primary and secondary external drivers are the same as described in Section 4.5.1, except for the initial value of the secondary driver. In this trajectory, greatly reducing the annual watershed NO_3-N export is still highly valued, but a slightly more conservative 30% reduction target has been imposed (Figure 4.10, center box). The optimal locations for *Miscanthus* (from the human perspective) are still in the high yield Salt Creek and Upper Sangamon tributaries, but overall, less *Miscanthus* is required to meet this goal (about 30% of the agricultural land

area), and the average annual profit is reduced by about 15% compared to the BAU period. At the outlet, the average annual NO_3-N export is 32% lower than BAU, and while the 7 day minimum flow for the entire 12 year period is reduced relative to the BAU period annually, only the very driest years are greatly affected (in this case, two). As before, the optimal locations for planting *Miscanthus* lie upstream of the same three lakes, reducing the inflow to each, but to a slightly lesser extent in Lake Decatur, since less of its drainage area has been converted to *Miscanthus* (Figure 4.10, center box).

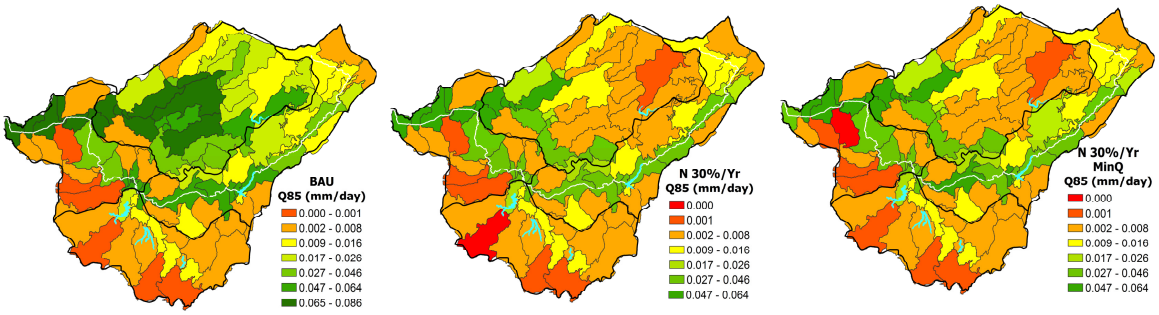

Figure 4.11: Spatial distribution of threshold low flows (Q_{85}) for three stages of Trajectory 2.

At the end of the second stage of Trajectory 2, the feedbacks from the hydrologic system (reduction in low flows) at the outlet may not be large enough to be noticeable to the human system, and 30% agricultural land conversion to biofuel crops may or may not be considered a major cause for concern, at least after only 12 years. Thus a plausible human system response in this case, based only on assessment at the outlet, might be to adopt a "wait-and-see" attitude, and reassess the situation when the current *Miscanthus* fields need to be replanted in another 12 years. However, if the hydrologic feedbacks at the sub-watershed scale are taken into account, a different picture emerges (Figure 4.11). The left panel in Figure 4.11 shows the threshold low flow for each sub-watershed under the BAU condition. Of interest is the heterogeneity in the threshold low flow (Q_{85}): Salt Creek, which is highly branched, has the highest low flows, while South Fork has the lowest, even after normalizing for drainage area. From the BAU condition to the 30% NO_3-N reduction biofuel development

plan, sub-watersheds most affected by *Miscanthus* show a 50 – 100% reduction in the Q_{85} low flows, with downstream sub-watersheds experiencing about 30% - 40% reduction, as shown by the red and dark orange colors in the insert map in Figure 4.10. The summer low flow period is when water stress is most likely to occur in this region, and one of the two affected water supply reservoirs already suffers water shortages during dry years under BAU conditions (Yaeger et al., 2013); biofuels development has exacerbated this problem (Figure 4.11, center panel). Another possible response from the human system could be to use flow constraints to protect vulnerable water supplies while at the same time attempting to maintain the level of water quality improvement gained over the previous 12 years. These flow constraints, which are based on the historical minimum monthly flows, are imposed upstream of Lake Springfield and Lake Decatur.

When minimum flow constraints are imposed within the watershed, the result is a shift in the locations of *Miscanthus* cultivation and a slight increase in total land area converted, since lower yield (and perhaps higher transportation cost) areas are utilized. With 32% of the agricultural land now converted to *Miscanthus*, annual average system profits increase very slightly, while the outlet NO_3-N export is actually reduced by 3%, due to 2% more *Miscanthus* in the watershed compared to the previous stage. The 7 day minimum flows at the outlet decrease due to expanded *Miscanthus* cultivation immediately upstream in the Lower Sangamon. While Lake Springfield inflows completely return to BAU conditions, Lake Decatur's do not, since some *Miscanthus* remains in the Upper Sangamon (Figure 4.11, right panel); however, inflow remains above the historical minimum.

4.6 Discussion

The results presented in the previous section show some of the multiple trade-offs that must be considered for decision making to reach a satisfactory balance between the competing needs for water of sufficient quality of agriculture, (bio)energy, and the environment, and which will affect the co-evolution of the coupled system. This section will further discuss the similarity and difference between the two projected trajectories, and illustrate these trade-offs in more detail. Additional discussion will also be provided on the feedbacks between

natural and human systems given that human society has now shown more concern for ecosystem integrity in watershed management.

In many ways, the two trajectories presented in Section 4.5 are very similar. Both trajectories operated under similar assumptions; the same primary external driver (the RFS2 mandate) was applied to each trajectory, and at all stages, reducing NO_3-N export was a secondary driver of human system decisions, since excessive riverine NO_3-N export has been, and still is, a serious problem in the Sangamon watershed. In both cases, the same high-yield *Miscanthus* regions of the watershed were chosen, at least at first, for conversion to dedicated biofuels crops. This is likely due to local climate, soil, and drainage properties contributing to the high biomass yield and this, coinciding with lower transportation traffic and proximity to ethanol demand zones in that region, caused the SoS model to choose those locations in all unconstrained trajectory stages. Because of these similarities, the same water users (municipalities, power plants, and point-source discharge permit holders) were affected in both trajectories. Furthermore, because many of these optimal *Miscanthus* locations were also smaller-order reaches, which derive a higher proportion of streamflow from hillslope drainage (Yaeger et al., 2013), summer low flows became increasingly intermittent in both cases. Overall, the similarities in the two trajectories can mainly be explained by or related to the hydrology and climate of this watershed.

Because of the similarities between the two trajectories, similar trade-offs were highlighted by the modeling framework in all cases. Of importance to sustainable biofuels development were the trade-offs between hydrologic properties (water quality and water quantity), between land uses (land for food crops and land for fuel crops), and between water users (municipalities, power plants, and point source permit holders). These trade-offs represent competing interests within each system, requiring that some satisfactory balance be found between them; how this balance was defined and achieved is where the two trajectories diverged, thus altering the future co-evolution of the coupled system.

In spite of these similarities, however, each trajectory followed a different path. The magnitude of the secondary external driver imposed in Stage 2 in each case led to differences in the resulting internal responses of the coupled system, causing the trajectories to diverge. This is shown in Figure 4.12, where the black line represents a nonlinear relationship between

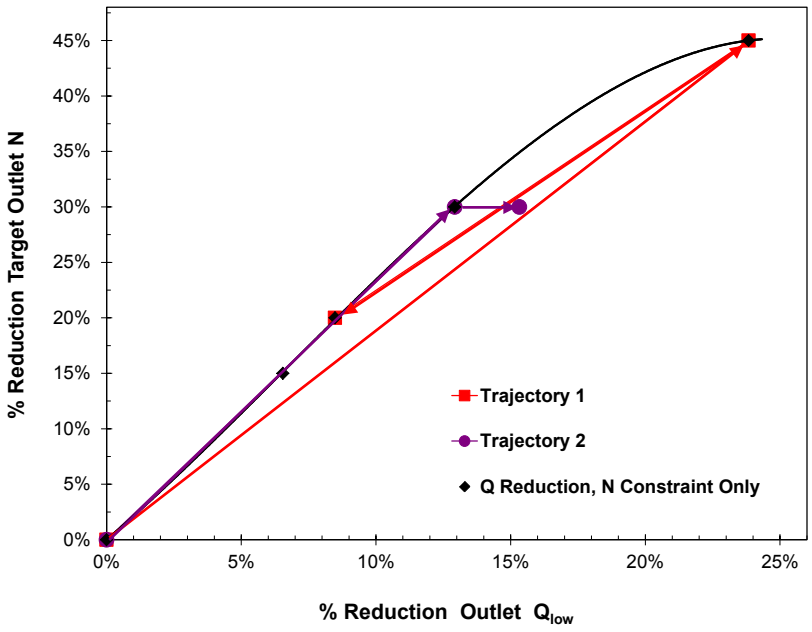

Figure 4.12: Alternate pathways for balancing water quality improvement (NO_3-N reduction at outlet) and water quantity damage (low flow reduction at outlet).

the reduction of low flow and NO_3-N at the outlet. When the magnitude of the secondary driver (the NO_3-N reduction target) is large (e.g. 45%), so much of the watershed is directly impacted by *Miscanthus* cultivation that even the outlet is seriously affected. Although there are many local feedbacks from the hydrologic system to which the human system could choose to respond, the cumulative effect is noticeable enough at the outlet that a watershed-scale response is a reasonable option. Thus, Trajectory 1 follows the relationship in Figure 4.12, finding a balance at a lower reduction target value (20% N reduction annually), with low flows reduced by about 8% at the outlet. Even though the local hydrologic feedbacks were not directly addressed, some of these (e.g. the municipal water supply reservoirs) were addressed indirectly by reducing the amount of *Miscanthus* overall. In contrast, when magnitude of the secondary driver is lower (e.g. 30%), although critical flows were reduced at the watershed outlet, the decrease was mainly noticeable only during the driest years, and thus might be ignored by the human system, particularly if there is no minimum flow requirement at that location. This is when the local feedbacks become important, and the human system may then need to prioritize the different water users, depending on how much water quality benefit

is desired. For the Trajectory 2, this meant leaving the NO_3-N reduction target unchanged at 30%, and instead choosing to protect the two municipal water supplies. Due to complex interactions between various subsystems of the biofuels development system and with the environment, Trajectory 2 diverges from the relationship in Figure 4.12, finding a different balance for the same NO_3-N reduction (30%) and a slightly worse low flow reduction of 15% at the outlet. In doing so, the water quality benefits gained for the watershed as a whole are balanced against water quantity impacts at the sub-watershed scale. Overall, the differences in the two trajectories can mainly be explained by the human system decisions and the unique mosaic of human adaptations and modifications within the watershed. That is, the scale of the human system response – global (entire watershed) vs. local (sub-watersheds) – combined with the history of past changes within the watershed control the trajectory of the coupled system.

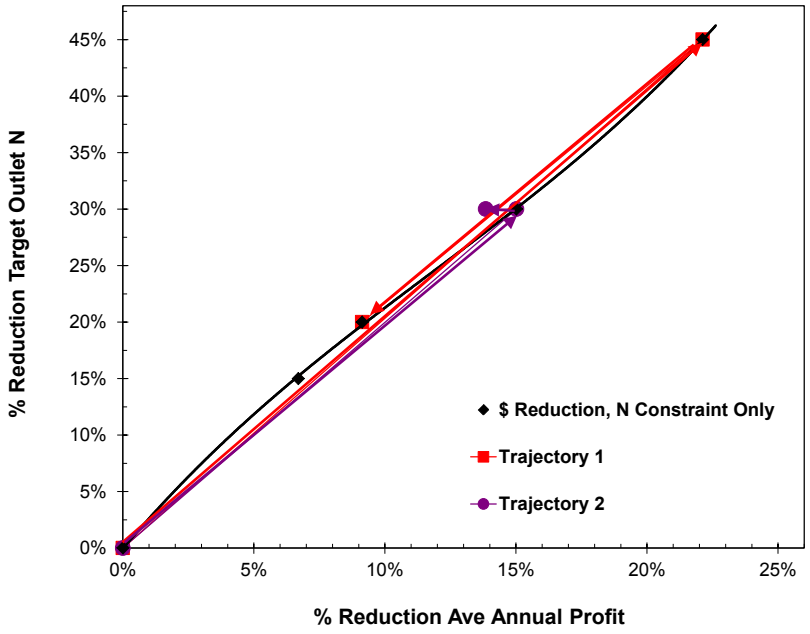

Figure 4.13: Alternate pathways for balancing water quality improvement (NO_3-N reduction at outlet) and water quantity damage (low flow reduction at outlet).

Analysis of the historical trajectory of this watershed in Section 4.1 had revealed a cross-scale disconnect between the large-scale water quality feedbacks from the hydrologic system and the human agricultural system impacts that are the source. Furthermore, human system

adaptations and technology have reduced human sensitivity to and awareness of local-scale feedbacks from the hydrologic system. However, within the integrated modeling framework presented here, a large-scale policy-driven human response to these large scale hydrologic feedbacks removed some of this disconnect by creating local hydrologic feedbacks that directly impacted local communities by noticeably reducing water quantity as well as the local watershed economic benefit (Figure 4.13). The NO_3-N reduction policy adopted in addition to the RFS2 mandate provided the desired water quality benefits in each case, but at a direct cost to the economic benefit. As with the water quantity reduction, the economic response is also nonlinear (Housh et al., 2013), and the resulting balance between water quality and profit found in each trajectory depended on the scale of the human system response. For Trajectory 1, watershed-scale response to hydrologic feedbacks (reducing the NO_3-N reduction target from 45% to 20% per year) allowed for a substantial recovery of profits, with the smaller target resulting in a 9% reduction in mean annual profits (Figure 4.13). In Trajectory 2, where the NO_3-N reduction target was initially 30%, the local response to hydrologic feedbacks (imposing flow constraints upstream of water supply reservoirs) resulted in a slightly smaller recovery of profits; since the NO_3-N reduction target remained 30%, the amount of *Miscanthus* grown in the watershed, and therefore the profit, changed very little (Figure 4.13).

The system trajectories chosen to demonstrate the application of this integrated modeling framework are only two examples of the many possible ways this system could co-evolve, depending on shifts in climate, development decisions, and response to hydrologic feedbacks. Future studies can generate more trajectories and analyze the various factors affecting the co-evolution of hydrologic processes and watershed management practices. Also, the current study mainly considered human water use needs, with ecological needs assumed to be met with sufficient minimum flow. If the needs of environmental users such as specific fish functional guilds and other aquatic ecosystems are considered as well as the known human water demands, additional locations within the watershed may require sufficient flow or flow characteristics, which would further change the spatial configuration of biofuels development. Thus, in addition to the competition and trade-offs already discussed, there exists, but is not considered in this study, a trade-off between ecological and human water users. While

the ecological side has been ignored in the past, leading to some ecological degradation, human system values have been changing. Under the 1972 Clean Water Act, the USEPA has been overseeing ecosystem monitoring and restoration, and ecosystem integrity is now being accounted for in watershed management institutions. This shows the human system response to ecological feedbacks as well as purely hydrological ones; thus, the modeling framework can, and should, be adapted to take account of these interactions as well, since they will likely further influence the future co-evolution of the coupled system. Because the ecological aspects are beyond the scope of this paper, we leave this implementation to future work. Lastly, a third limitation of this framework which must be acknowledged is the uncertainty inherent in the output of both models. However, the trends produced in the interactive use of these models within the framework are reasonable, and since our goal is predictive insight, not prediction in the traditional deterministic sense, there is still value in this work. Furthermore, a better understanding of the relevant processes and interactions within the system of interest gained from analyses using this framework may also be applied toward improving the modeling components.

4.7 Conclusions

This paper has presented a new quantitative framework with which to analyze the dynamic interactions within coupled human and nature systems, and its implementation in an intensively-managed Midwestern US agricultural watershed in the context of biofuel development. The framework coupled a System of Systems (SoS) optimization model for human decisions with respect to biofuel development and a watershed model for hydrologic and water quality feedbacks in an interactive way, capturing the co-evolutionary dynamics of the coupled system as it spiraled forward in time and produced likely trajectories of future change. Because of the connection between agricultural development in the study region and the problem of hypoxia in the Gulf of Mexico as reflected in the regions known past historical trajectory, the complexity of human system responses to the imposed biofuel mandate was constrained by a secondary driver consisting of an annual watershed NO_3-N reduction target. Imposition of the external biofuel mandate combined with different secondary environmental

targets resulted in nonlinear and multi-scale responses of both the human and hydrologic systems, including multiple trade-offs, and impacting the future co-evolution of the system in complex, heterogeneous ways.

This integrated modeling tool enables the generation of a whole spectrum of future trajectories, depending on human adaptations and changes to human values and norms. In this paper, two such trajectories were explored, each developing progressively in three decadal stages, allowing for variable human feedbacks in response to the biofuel and nitrate outcomes produced in the preceding stage. Both trajectories had the same cellulosic ethanol mandate, but differed in the initial secondary NO_3-N reduction target imposed: 45% for Trajectory 1 and 30% for Trajectory 2. In both cases the beneficial, targeted reduction in NO_3-N also came with significant detrimental reduction in land area for food crops, annual profits, and summer low flows, all of which were smaller in scope in the latter case. In the case of Trajectory 1, the detrimental reductions were noticeable at the outlet (the watershed scale), and prompted a subsequent adaptive response to these feedbacks at the same scale, with the humans reducing the NO_3-N target at the outlet to only 20% to balance out the benefits and detrimental effects. In the case of Trajectory 2, however, the control shifted to water deficits at the local (subwatershed) scale with respect to city water supplies. The trajectories can be deemed as internal responses of the coupled human-hydrologic system to the imposed biofuel mandate and nitrate constraint, and these examples show that, depending on the magnitude of the external drivers and constraints, they manifest differently in the resulting dynamic changes in land use, average annual profits, and both water quality and quantity, thus showing the reciprocal feedbacks between human and natural systems.

The trajectories explored here showed that the external drivers were propagated through the coupled system in complex ways at different scales, manifesting differently in different places. Because both model components of the adopted framework were distributed, the heterogeneity of both the human and hydrological systems was able to be captured. The interdependencies between the various subsystems in the SoS model created outcomes for the optimal biofuels land use development which then propagated through the hydrologic model from upstream to downstream, creating further complexity in the response. Furthermore, the coupled co-evolutionary modeling framework was designed to mimic the progressive

human system responses to the hydrologic feedbacks, thus allowing for "course corrections" along the way. Because of this, the process of developing the trajectories provided valuable predictive insights into the dynamics of the coupled system. For the case study system, the result of using the cellulosic ethanol mandate to reduce NO_3-N export to the Gulf was not a simple trade-off of water quantity for water quality, but instead involved multiple trade-offs affecting multiple subsystems and stakeholders, notably the trade-off between land for food crops and land for fuel crops, as well as the trade-offs between different water users within the watershed, such as cities, power plants, and discharge permit holders. The richness of the results possible from the use of the adopted framework also reflects the interaction between the external drivers and the internal mosaic of previous human modifications to the landscape and hydrology over the watersheds entire co-evolutionary history, which is another key strength of the integrated modeling framework.

Acknowledgments

The data for this paper is publicly available at the locations listed in Table 4.1. We also thank M. B. David for providing the water quality data used to calibrate the SWAT model in this study. The research presented here was done as part of the NSF-funded project "Interdependence, Resilience, and Sustainability of Infrastructures for Biofuel Development" (NSF grant EFRI-083598, X. Cai, PI).

Chapter 5

Water for food, energy, and the environment: predictive insights from comparative analysis in an integrated modeling framework

Abstract

There are serious challenges to finding the balance between water for food, energy, and the environment, notably the complexity of the interactions between the human and hydrologic systems while both are continuously undergoing change. These challenges necessitate an improvement in predictive ability to capture the dynamics of water resources in the long term. To identify, describe, and understand the interactions and feedbacks between the human and hydrologic system the two systems need to be treated as a linked, co-evolving system. This study expands on previous work, utilizing the insights gained from a comparative analysis of historical trajectories of two intensively-managed agricultural watersheds and applying them in previously-developed integrated modeling framework to comparatively explore future response trajectories. Results showed that historical co-evolution of humans with the natural system results in different spatial patterns of modification that influence internal responses to external driving forces, highlighting the importance of location and history in determining the direction future trajectories will take. Therefore, to make more informed water resource decisions, in this region at least, watershed scale studies are necessary because each watershed will have a unique mosaic of change, even if they are geographically close and climatically similar. Large, more global scale studies would miss these local differences, thus policy decisions based on the resulting information could miss potentially serious local impacts.

5.1 Introduction

Sustainable management of finite water resources requires balancing the changing water needs of society and ecosystems under increasingly uncertain conditions (Wallace et al., 2003). A major challenge facing hydrology today is that the decisions necessary to accomplish this require long term predictions about the future state of a changing system under nonstationary conditions (Milly et al., 2008; Wagener et al., 2010), a task for which traditional deterministic methods may not be adequate (Peel and Blöschl, 2011; Sivapalan et al., 2011b). Recent studies proposing new frameworks within which to mediate this competition between human and environmental water needs have emphasized watersheds as co-evolving (Thompson et al., 2013; Harman and Troch, 2014) coupled human and nature systems (CHANS) (Liu et al., 2007b). A specific example of this motivates the present study, which continues work begun in Yaeger et al. (2014).

Here, the general problem of managing water resources and land use change in a biofuels development context is complicated by the geographic extent that is affected (most of the Midwestern US, including the Upper Mississippi River Basin), and the long history of human modifications to the region (Yaeger et al., 2013). While many studies have been done to predict the possible hydrologic responses to proposed biofuels development, the spatial scales of the analyses vary widely. Because some of the hydrologic effects of land use change in the Midwest actually manifest in the Gulf of Mexico at the mouth of the Mississippi, some studies encompass the entire Mississippi River Basin (Donner and Kucharik, 2008). Others focus on the regions undergoing the most change, ranging in scale from the Midwestern states (Vanloocke et al., 2010), to the Upper Mississippi River basin (Secchi et al., 2011; Demissie et al., 2012; Wu et al., 2012), to individual river basins within this region, (Wu and Liu, 2012), and then to smaller watersheds (Ng et al., 2010; Housh et al., 2014). However, it is not clear what is the appropriate scale at which to study CHANS dynamics and interactions.

In presenting the idea of earth surface systems as evolving landscapes, Phillips (2002) proposed that such systems are governed by global physical laws acting at individual locations over time to produce unique response trajectories to changes. Given the complexity of these systems, internal responses to external driving forces can be highly nonlinear (Thomas,

2001; Phillips, 2003), and deterministic long term prediction of future states becomes nearly impossible (Phillips, 2006). When the human system is considered as co-evolving along with the hydrologic system, this importance of geography (location) and history increases since humans tend to adapt and modify the environment to suit their needs (Phillips, 2001; DeFries et al., 2004; Liu et al., 2007a; Gordon et al., 2008). This was highlighted in the previous study Yaeger et al. (2014), where even in a fairly homogeneous watershed, the spatial mosaic of historical human adaptations to the climate and landscape created nonlinearity in the internal responses of the coupled system to external driving forces applied to it. These results indicate that the watershed scale is more suitable to better understanding CHANS interactions, since larger scales may miss or average out location-specific modifications.

Because of this importance of location and history in the co-evolution of CHANS, more can be learned about underlying processes and interactions by comparative study of different locations (Harman and Troch, 2014). This can then be used to constrain possibilities of future trajectories (Phillips, 2007) in a co-evolutionary modeling framework and test hypotheses about behavior trends thus revealed. In this way, predictive insights can be developed that could be used to guide management decisions (Thompson et al., 2013). For the present study, comparative analysis of historical trajectories and present states of two adjacent intensively managed agricultural watersheds in the Midwestern US revealed how the unique mosaics of both human and natural changes over time resulted in very different hydrologic behavior in the two watersheds (Yaeger et al., 2013). Areas with extensive tile (subsurface) drainage and row-crop agriculture were associated with a higher baseflow index (BI), shorter residence times, and greater NO_3-N (nutrient) export. Areas with higher topographical relief required less tile drainage, and this, along with more heterogeneous land cover (forest, pasture, and row crop agriculture) resulted in lower BI, longer residence times, and less NO_3-N export. A second important difference was in the size and operation of reservoirs in each watershed: in the more tile-drained watershed, reservoirs were small, used for water supply or cooling ponds, and with little control on the flow regime, while in the second watershed, large flood control structures modified the flow regime. Because of the much longer retention times of these reservoirs compared to the smaller ones used for water supply, significant NO_3-N removal takes place (McIsaac and Hu, 2004; David et al., 2006), providing a secondary water

quality benefit in addition to flood control. These differences in the human modifications to the hydrology of each watershed influenced both water quality and quantity and this resulted in a nonlinear internal response to imposition of external drivers (bioethanol mandate and NO_3-N reduction targets) on the tile-drained, row-cropped watershed.

The present study expands on this work, using the integrated modeling framework previously developed to comparatively explore future response trajectories of the same two intensively-managed agricultural watersheds just described. We hypothesize that comparing the differences in the propagation of the same external drivers (climate, biofuel mandates, and nutrient reduction targets) through both watersheds will further our understanding of the underlying processes. The aim of this paper is to comparatively analyze both case study CHANS using the previously-developed integrated modeling framework by projecting their co-evolution into the future in stages in order to develop predictive insights into how external drivers propagate through the coupled system, and in doing so identify some of the multiple trade-offs necessary at the watershed scale for sustainable water resource management in a biofuels development context.

The aim of this paper is to further our understanding of the underlying processes and develop predictive insights that can guide sustainable water resource management. This paper is outlined as follows: in Section 5.2 a brief hydrologic overview of the case study catchments is presented, and the integrated modeling framework is outlined. Section 5.3 comparatively presents the modeling results, including the results from the previous study for contrast while Section 5.4 discusses the similarities and differences between watershed trajectories. The last section summarizes the main findings.

5.2 Methods

5.2.1 Case study watersheds

The study area consists of two watersheds in central IL (Figure 5.1), the Sangamon River in the north, with a drainage area of 14,000 km^2, and to the south is the Kaskaskia River watershed, a roughly 15,000 km^2 tributary of the Mississippi River. Although each watershed

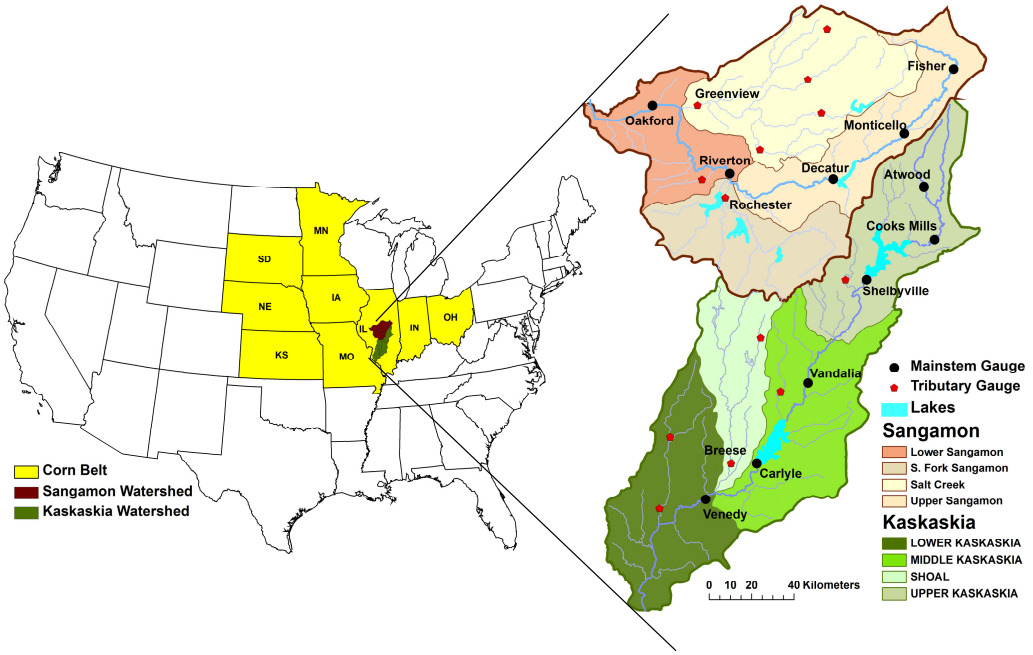

Figure 5.1: Locator map of the two watersheds

consists of four main hydrologic units (HUCs), the hydrologic connectivity differs between them. While the Kaskaskia River mainstem is divided into an Upper, Middle, and Lower HUC, with one large tributary, Shoal Creek, emptying into the Lower Kaskaskia, the Sangamon watershed has only an Upper and Lower main-stem and two hydrologically separate tributaries – Salt Creek in the north and South Fork in the south, with confluences in the Lower Sangamon. While both watersheds are fairly flat, there is more topographical relief in Kaskaskia compared to Sangamon, which mainly consists of a flat upland plateau and a flat lowland region. A topographical divide in the northeast ridge extending from the Salt Creek and Upper Sangamon into the Upper Kaskaskia separates the upper portion, which shares soil, vegetation, and drainage characteristics with Sangamon, from the rest of the Kaskaskia watershed, which is marked by gentle slopes to deeper valley bottoms. These valley slopes provide enough topographical relief that the Middle and Lower Kaskaskia are more well-drained naturally (Prince, 2008) than the Sangamon and the Upper Kaskaskia.

The climate in this region is humid and temperate, with an average of about 1000 mm of precipitation annually, although severe, extended droughts occur periodically (Woodhouse and Overpeck, 1998). Precipitation is not distributed uniformly, but is highly heterogeneous,

both spatially and temporally. At the annual scale, this variation is seen both between years and between locations, with a regional north-south gradient to the annual precipitation, where the southern portions of Kaskaskia on average receive more rainfall annually than the northern parts of Sangamon (Yaeger et al., 2013). Precipitation also varies seasonally within the year, with April and May generally being the wettest months and September and October the driest. Summer precipitation tends to be localized, convective, and highly intense, and this spatial heterogeneity at small scales interacting with and filtering through the historical modifications to the system, increases the complexity of the catchment hydrologic response.

Both the Sangamon and the Upper Kaskaskia require extensive tile drainage for crops to be successfully grown, due to the combination of flat watershed topography and deep rich soils underlain by a hardpan subsoil layer, while the Middle and Lower Kaskaskia require little tile drainage. The land use in both watersheds is mainly annual row crop agriculture (90% Sangamon and 70% Kaskaskia), with the rest mainly forest in Kaskaskia and some urban and riparian forest in Sangamon. In Kaskaskia, there are two large flood control reservoirs; Lake Shelbyville in the Upper Kaskaskia and Lake Carlyle in the Middle Kaskaskia (Figure 5.1). In Sangamon are several small, uncontrolled, inline reservoirs, three of which serve as municipal water supplies and two as cooling ponds for power plants. Of interest is Lake Decatur, which supplies water for the city of Decatur (Figure 5.1); due to its location in the Upper Sangamon, which experiences more variable streamflow than the lower reaches, it is more vulnerable to extended drought periods (Yaeger et al., 2013).

5.2.2 Integrated Modeling Framework

The integrated modeling framework employs a system of systems optimization model to emulate human development decisions which are then incorporated into a watershed model to estimate the resulting hydrologic impacts. The two models are run interactively in 10 – 20 year stages to simulate the co-evolution of coupled human-nature systems, such that reciprocal feedbacks between hydrologic processes and human decisions (i.e., human impacts on water quality and critical low flows and hydrologic impacts on human decisions on land and water use) can be assessed. The primary anthropogenic external driver imposed on

these two systems is a bioethanol mandate requiring a minimum amount of both cellulosic and total ethanol to be produced. Because improvement to water quality has been a priority in this region, a secondary external driver constraining NO_3-N export is also imposed, but is subject to change depending on internal feedbacks. Possible future trajectories developed in this way are presented in three stages: first, the optimized current conditions, or "business as usual" (BAU); second, the internal response to imposition of the primary external driver and an initial secondary driver; and third a response to feedbacks from the hydrologic system due to human impacts in the second stage, subject to the primary external driver. Further details of the framework can be found in Yaeger et al. (2014).

The System of Systems (SoS) Biofuels model (Housh et al., 2013) is a multiobjective optimization model that includes all subsystems in the biofuel development system. The overall objective is to maximize the net economic profit for the average optimal year, over an 11 year rolling horizon. Because of the unique hydrologic infrastructure in the Kaskaskia watershed, cascading reservoirs have been added to this version of the model to represent Lake Shelbyville in the Upper Kaskaskia, and Lake Carlyle in the Middle Kaskaskia. The release rules for each reservoir are represented by an inflow/outflow relation that follows the seasonal pattern of release as given by the US Army Corps of Engineers (Stemler, 2011). An average denitrification rate of 50% was applied to both reservoirs in the SoS model, based on the values obtained for Lake Shelbyville (David et al., 2006). Furthermore, there are minimum flow constraints already in place in this watershed; Lake Shelbyville a 10 ft^3/s minimum release rate, and Lake Carlyle a 50 ft^3/s minimum. These values were converted to million m^3/mo and applied in the SoS model at the outflow only for Shelbyville, and at the outflow to the watershed outlet for Carlyle, due navigation requirements for the Lower Kaskaskia.

The Soil and Water Assessment Tool (SWAT) version 2005 developed by the US Department of Agriculture (USDA) Agricultural Research Service (ARS) (Arnold and Fohrer, 2005; Gassman et al., 2007) was selected as the watershed model because it is well-documented, the required input data is readily available (Table 5.1), and it has been used extensively in similar biofuels development applications in the Midwest (Schilling et al., 2008; Ng et al., 2010; Secchi et al., 2011; Demissie et al., 2012; Wu and Liu, 2012; Wu et al., 2012), as well as

the Kaskaskia specifically (Chiang et al., 2014). Of these, Ng et al. (2010) developed SWAT crop growth parameters for *Miscanthus* which we have used in the simulations presented in this paper.

Table 5.1: Description of and sources for data used in this study

Data Description		**Data Source**
30-m Land cover, 2001 30-m digital elevation map (DEM)	USGS[1] National Map	http://nationalmap.gov/viewer.html
STATSGO Soils map, by state	USDA NRCS[2]	http://soildatamart.nrcs.usda.gov/USDGSM.aspx
Daily precipitation	NOAA[3] NCDC	http://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn
Nitrate-N concentrations (downstream calibration)	US EPA STORET	http://www.epa.gov/storpubl/legacy/query.htm
Daily streamflow (calibration)	USGS NWIS	http://maps.waterdata.usgs.gov/ mapper/index.html
Annual Crop Yields, by state and county (calibration)	USDA NASS	http://quickstats.nass.usda.gov/

[1]US Geological Survey, [2]Natural Resources Conservation Service, [3]National Oceanic and Atmospheric Administration

For the comparative modeling analysis, SWAT models of both watersheds were developed. Details of the SWAT model calibration for the Sangamon watershed can be found in Yaeger et al. (2014); this paper will focus on the development of the SWAT model of the Kaskaksia watershed. Current local agricultural practices were modeled after the land management scheme presented in Hu et al. (2007), while ranges for other parameter values were obtained from the literature cited above. In total, the watershed was delineated into 94 subwatersheds, ranging in area from 10 km^2 to nearly 100 km^2. Following Ng et al. (2010), since over 70% of the watershed is agricultural, multiple HRUs per subwatershed were not implemented. In light of recent studies of this watershed, both data-based (Yaeger et al., 2013) and modeling with SWAT (Chiang et al., 2014) this assumption may not be valid in this case. Only the largest reservoir, Lake Carlyle, was incorporated into the SWAT model.

The model was calibrated for a total of 9 years (1994 to 1998 for the downstream, and 1998 to 2002 for the upstream). This period was chosen not only because it incorporated both dry and wet years, but also because it was necessary to ensure that there was enough

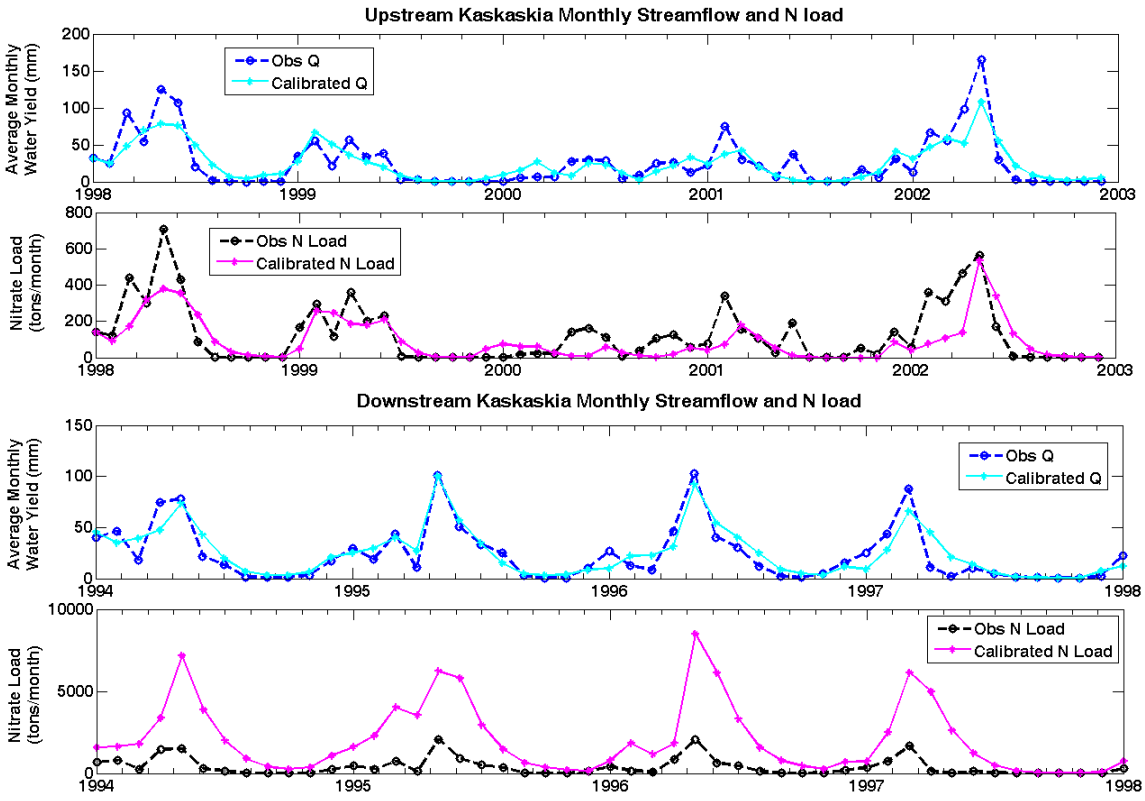

Figure 5.2: Monthly calibration results for Kaskaskia upstream and downstream.

NO_3-N data for validation. For the downstream, only legacy (1998 and before) EPA data was available, while more recent data was available for the upstream. The calibration of the finished model was carried out according to the procedure outlined in Neitsch et al. (2010) for daily streamflow at four points along the mainstem and one at the outlet of each of the main tributary, and monthly NO_3-N load upstream in Upper Kaskaskia and downstream in Lower Kaskaskia below Lake Carlyle (Figure 5.2). The Nash Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), root mean squared error standard deviation ratio (RSR), and percent bias (PBIAS) were used to determine goodness of fit between modeled and observed data as outlined in Moriasi et al. (2007). The NSE for monthly discharge, ranging from 0.55 to 0.82, were mostly within literature recommendations (Moriasi et al., 2007), but those for daily discharge (0.14 0.56) were good for the mainstem locations, but very poor for the non-tiled tributary in the downstream. The NSE for monthly NO_3-N was good for the upstream location (0.51), but very poor (<0) for the downstream. Additionally, RSR was <1 for all gauges, and PBIAS was between 3 and 7% (showing overestimation) for gauges

above the main reservoir and -7% to -19% (showing underestimation) below it and on the main tributary.

In general, model performance improved with increasing drainage area, but the model does not fully capture the heterogeneity seen in the data Yaeger et al. (2013). This is likely due in large part to simplifying assumptions made when developing the model. While tile drains, one reservoir, and some of the soil heterogeneity are included in the model, the land use heterogeneity in the Middle and Lower Kaskaskia, where much of the watershed's 30% forest cover is located, is not. This explains in part the overestimation of NO_3-N at the downstream, since this area is modeled as row crop agriculture, and not mixed forest and agriculture.

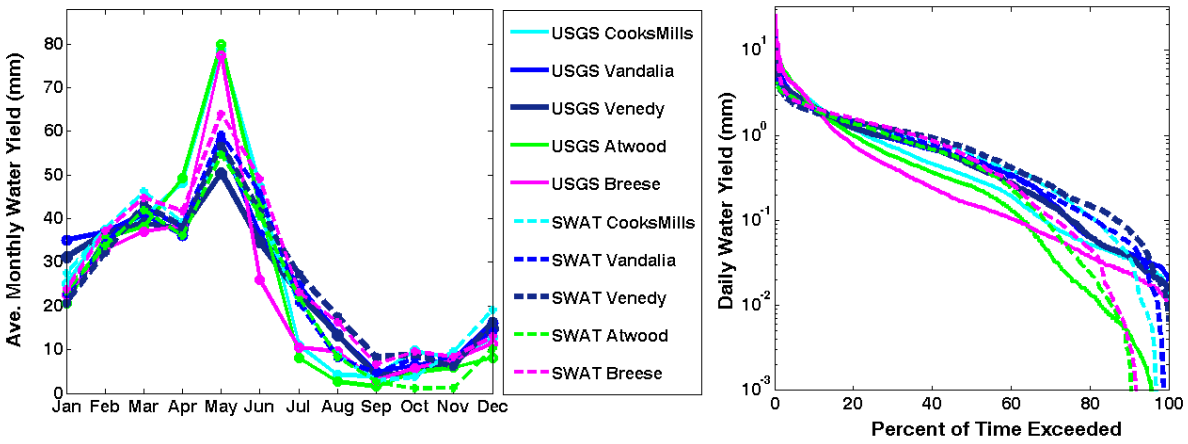

Figure 5.3: Simulated (dashed lines) and observed (solid lines) monthly regime curve (left) and daily flow duration curve (right) for 6 calibration/validation locations in the watershed.

Because low flows were of interest in this particular case study, the model was calibrated with care to capturing the daily flow regime. In addition to fitting the daily hydrograph, the model output was also checked against the monthly regime curve (RC) and daily flow duration curve (FDC) to ensure that the model could adequately represent the low flow patterns needed for the low flow analysis (Figure 5.3, left and right panels, respectively). These figures show the model tends to overestimate early summer flows and underestimate the spring flood peak, but overall, the general trends are captured. The three distinct flow regime patterns seen in the data (Yaeger et al., 2013) are not seen in the modeled RCs because some of the necessary heterogeneity was missing from the model. This is also seen

in the FDC, where average flows are overestimated, but low flows are both over- and underestimated. This shows the model development may be too simple for the purposes of this study, since it entails capturing fine scale heterogeneity

5.3 Results

The trajectories presented in the following section are based on the premise that the human system highly values reduction of NO_3-N export to the Gulf of Mexico, thus imposing a secondary external driver on the coupled system. The initial state in both cases was an optimized land use solution provided by the SoS model for the current cropping rotation and corn ethanol production, hereafter referred to as "business as usual" (BAU); the state of the hydrologic system at the watershed outlet under an 11 year period of these conditions provides a baseline against which the NO_3-N reduction can be calculated. At the end of this time, the cellulosic ethanol mandate is imposed, along with an initial NO_3-N reduction target, so that subsequent stages of the trajectories form consecutive impact/feedback loops of internal responses to these external drivers. Results from Kaskaskia are presented along with those of Sangamon (Yaeger et al., 2014) for comparative analysis.

5.3.1 Trajectory 1: Reduce annual watershed NO_3-N export by 45%

The initial condition of the coupled system is shown in Figures 5.4 and 5.5 (left box). Land use in both watersheds is mainly row crop corn/soy rotation, with Sangamon having a larger proportion of agricultural land than Kaskaskia. Nitrate-N export from both watersheds is fairly high, but Kaskaskia, with less tile drainage and two large reservoirs, exports less than Sangamon. Because this region has been highlighted as the source of the majority of N export to the Gulf, we assume the human system will use the water quality benefits of the dedicated biofuels crop *Miscanthus* to reduce annual export of NO_3-N export by the EPA-recommended 45% in both watersheds while also producing the mandated amount of cellulosic ethanol (Figures 5.4 and 5.5 center box). In Sangamon, this required about half of the agricultural land area to be converted to *Miscanthus*, with the majority of this

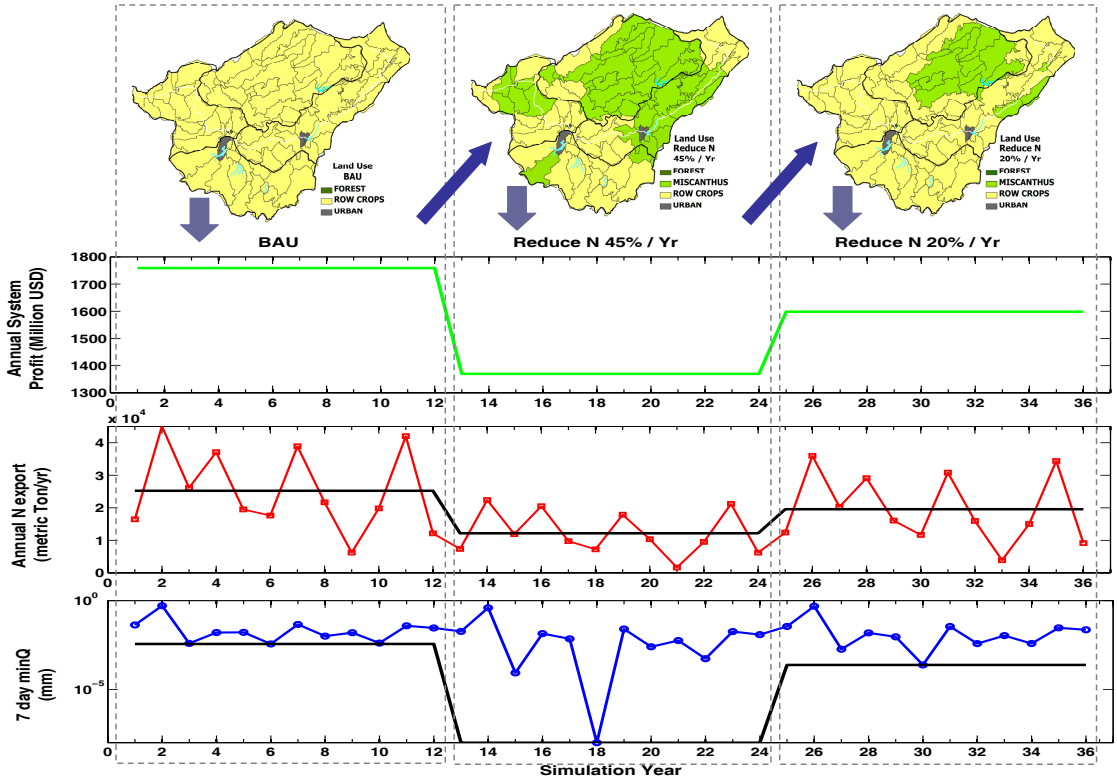

Figure 5.4: Sangamon Trajectory 1, where the initial goal is to meet EPA recommended 45% reduction in annual N export from the watershed (reproduced from Yaeger, et al., 2014). Note: land use maps, annual system profit, annual watershed N export, and annual 7-day minQ at outlet, are shown top to bottom, respectively; straight lines represent mean annual values for profit and N, and the minimum value for 7-day minQ.

conversion occurring in the high-yield tributaries Salt Creek and Upper Sangamon (Figure 4). In Kaskaskia, however, only 37% of the agricultural land is required to achieve this reduction (Figure 5.5), and this is clustered in the downstream end of the watershed. Furthermore, in Sangamon, this decreases the average annual system profit by 22%, while in Kaskaskia the system profit hardly changes.

The impacts of this land use decision on hydrology are mixed; while the water quality benefits are large in both watersheds, with total NO_3-N export reduced over 47% annually, both watersheds show a large drop in the lowest flows at the outlet, as expressed by the 7 day minimum flow, especially during the driest years. Within the Sangamon watershed, inflow to two of the municipal water supply lakes and one of the cooling reservoirs is reduced due to extensive *Miscanthus* planting upstream. Within the Kaskaskia watershed, the mainstem

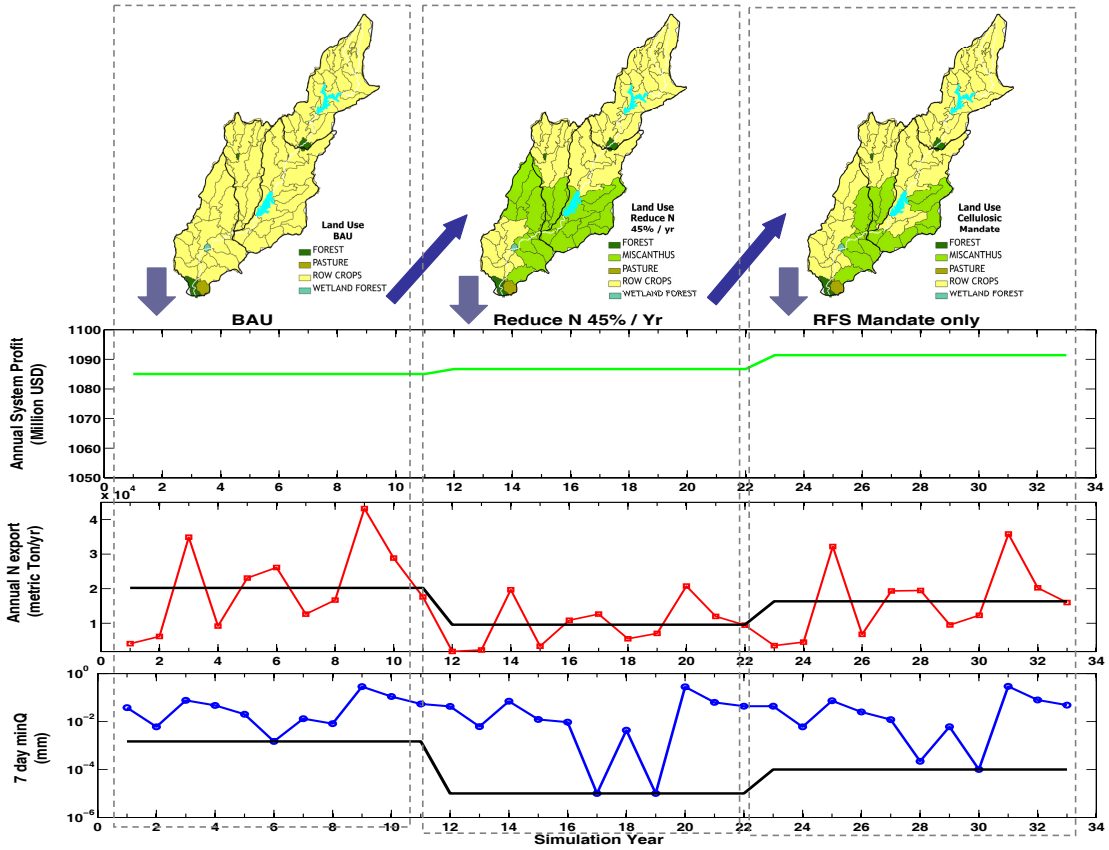

Figure 5.5: Kaskaskia Trajectory 1, where the initial goal is to meet EPA recommended 45% reduction in annual N export from the watershed. (Note: land use maps, annual system profit, annual watershed N export, and annual 7-day minQ at outlet, are shown top to bottom, respectively; straight lines represent mean annual values for profit and N, and the minimum value for 7-day minQ.)

inflow to the largest reservoir is unaffected, since *Miscanthus* is planted in the surrounding subwatersheds, the majority of which do not drain into the reservoir, as well as downstream of the reservoir. Thus in Kaskaskia, it is mainly the outlet that is most affected. In both watersheds, the outlet streamflow is greatly reduced, especially the low flows, although in Sangamon, there are significant local impacts as well.

Given the negative impacts on the human system (loss of agricultural land and reduction in profits, at least in Sangamon) and the mixed feedbacks from the hydrologic system (reduction in both NO_3-N export and low flows at the outlet and within the watershed in Sangamon), one plausible response might be to relax the NO_3-N reduction requirement in an attempt to find some balance between competing land uses (food or fuel) and competing water needs

(quality and quantity), as illustrated in Figure 5.5 (right box), where the NO_3-N export reduction requirement is now only 20% of the BAU export. For Kaskaskia, the response was also to the reduction in lows flows at the outlet, but here, there is more importance attached to it, since the Lower Kaskaskia near the outlet must have sufficient flow for barge traffic. For comparative purposes, the same decision was made for Kaskaskia as for Sangamon to reduce NO_3-N export to 20% of BAU values.

The new NO_3-N reduction target allowed some land to revert back to row crop agriculture. In Sangamon, the remaining 20% under *Miscanthus* cultivation provided a 22% NO_3-N reduction; the average annual system profit increased to within 10% of BAU levels. Although the response was to the reduced flow at the outlet, due to the various subsystem interdependencies considered by the SoS model, the new optimal locations for *Miscanthus* do not affect the water supply reservoirs as much, although the power plant is still affected. In Kaskaskia, however, 25% of the agricultural land remained under *Miscanthus*, providing well over the 20% annual NO_3-N reduction. Again, the average system profit hardly changed, increasing slightly by less than 1%. On closer inspection of the solution, it was found that the maximum NO_3-N constraint was non-binding, that is, the 19% NO_3-N reduction would have been obtained by simply meeting the mandate. Because the *Miscanthus* is still clustered in the downstream region, with the majority not connected to Lake Carlyle, the low flows do not recover as much as they did in Sangamon, although the minimum 7day minimum flow for the time period has increased.

5.3.2 Trajectory 2: Reduce annual NO_3-N export by 30%

In the second system trajectory, both the initial conditions and the primary external driver are the same as described in the previous section, except for the initial value of the secondary driver. While the goal is still to greatly reduce the annual watershed NO_3-N export, a slightly more reasonable 30% reduction has been chosen (Figures 5.6 and 5.7, center box). In Sangamon, *Miscanthus* is still clustered in the high yield Salt Creek and Upper Sangamon tributaries, but now only about 30% of the agricultural land area has been converted, reducing the annual NO_3-N export by about 32% and the average annual system profit by

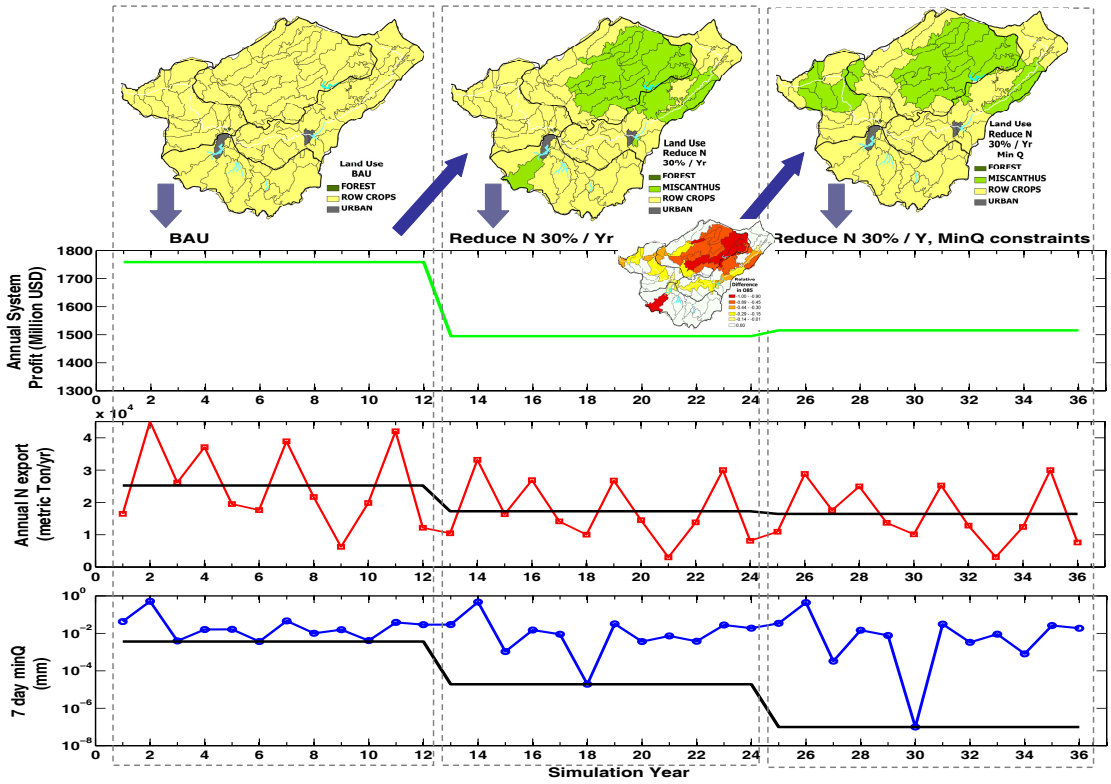

Figure 5.6: Sangamon Trajectory 2, where the initial goal is to reduce watershed NO_3-N export by 30%; the insert shows the hydrologic feedback given as relative change in Q_{85} flow (from the flow duration curve). Note: land use maps, annual system profit, annual watershed N export, and annual 7-day minQ at outlet, are shown top to bottom, respectively; straight lines represent mean annual values for profit and N, and the minimum value for 7-day minQ.

about 15%. In Kaskaskia, however, the 30% reduction goal is still already met and exceeded simply by meeting the cellulosic mandate alone, with 25% of the agricultural land converted to *Miscanthus*. The average annual system profit increases by a tiny fraction.

The hydrologic results are mixed, both within and between watersheds. In both watersheds water quality has been improved by a reduction in annual NO_3-N export of over 30%. In Kaskaskia, because of the optimal locations chosen for *Miscanthus*, the 7 day minimum flows at the outlet are affected worse than they are in Sangamon, where the 7 day minimum flow for the entire period is reduced relative to the BAU period, annually, only the very driest years are greatly affected. The optimal locations for planting *Miscanthus* in Sangamon lie upstream of the same three lakes, reducing the inflow to each, but to a slightly lesser extent

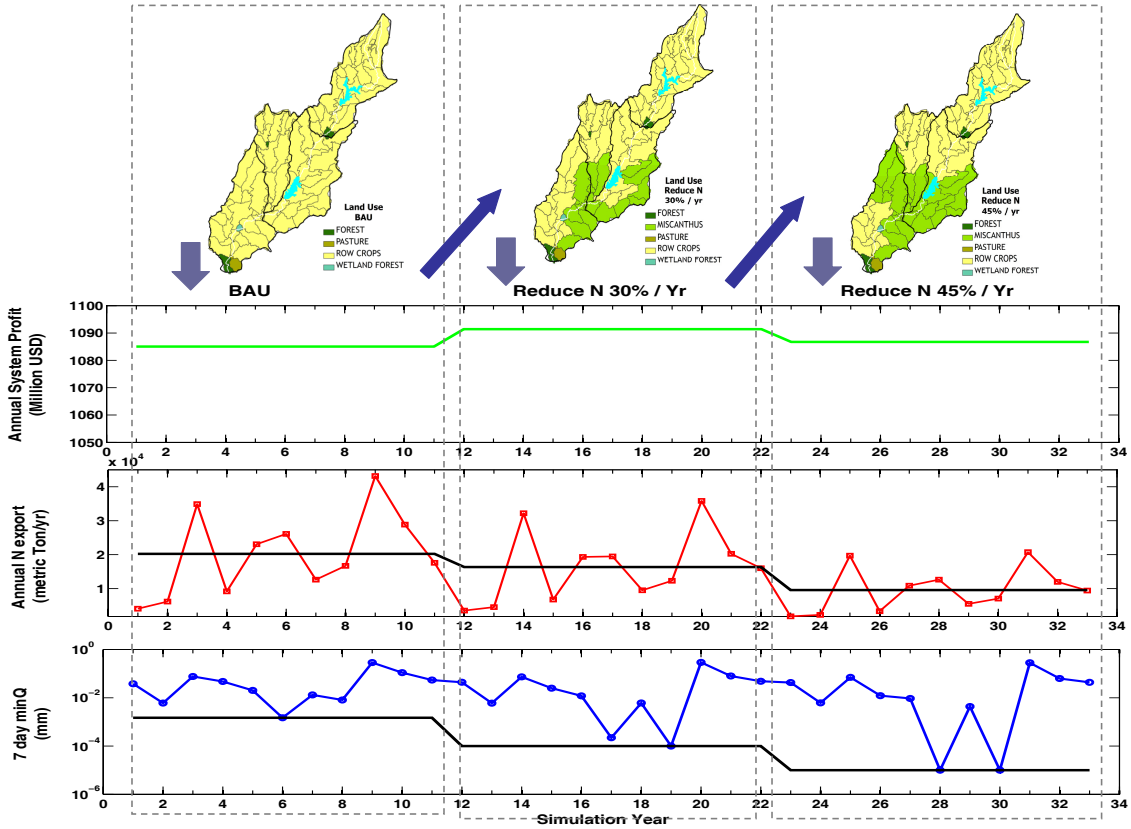

Figure 5.7: Kaskaskia Trajectory 2, where the initial goal is to meet EPA recommended 45% reduction in annual N export from the watershed. Note: land use maps, annual system profit, annual watershed N export, and annual 7-day minQ at outlet, are shown top to bottom, respectively; straight lines represent mean annual values for profit and N, and the minimum value for 7-day minQ.

in Lake Decatur, since less of its drainage area has been converted to *Miscanthus*.

At the end of the second stage of Trajectory 2, the feedbacks from the hydrologic system in both watersheds (reduction in low flows) at the outlet may not be large enough for the human system to take notice, and 25% – 30% agricultural land conversion to biofuel crops may or may not be considered a major cause for concern, at least after only 11-12 years. Thus a plausible response of the human system in this case, based only on assessment at the outlet, may be that everything is fine and do nothing, or wait until the current *Miscanthus* fields need to be replanted and reassess the situation. For Kaskaskia, since the impacts from the 30% NO_3-N reduction weren't too detrimental, and precipitation has been increasing over the past 50 years, the human system response may be to increase *Miscanthus* since

the watershed (and climate) appears to handle it. However, in Sangamon, if the hydrologic feedbacks at the subwatershed scale are taken into account, the local impacts on the water supply reservoirs and even the cooling ponds becomes evident. In this case, the human system would impose flow constraints upstream of vulnerable locations. For this trajectory, the constraints are only placed above the water supply reservoirs.

For Kaskaskia, increasing the NO_3-N reduction target results in even lower flows at the outlet which would affect the navigability of the river during very dry years. Interestingly, NO_3-N reductions using *Miscanthus* increase the profits slightly in Kaskaskia, unlike Sangamon, but only to a point, after which they decline. In Sangamon, when minimum flow constraints are imposed within the watershed, *Miscanthus* cultivation shifts into Salt Creek and into less-optimal regions downstream, resulting a slight increase in total land area converted to *Miscanthus*. The 7 day minimum flows at the outlet decrease due to expanded *Miscanthus* cultivation immediately upstream from the outlet. Inflow to the two water supply reservoirs remains above the historical minimum.

5.4 Discussion

From previous work (Yaeger et al., 2014) on the Sangamon watershed, N export has historically been a problem, due in large part to extensive tile drainage combined with intensive annual row crop cultivation. Tile drains increase baseflow and reduce soil residence times, resulting in less natural removal of NO_3-N from soil moisture before it reaches stream reaches. To meet the EPA 45% N reduction goal, nearly half of the cropland is converted to *Miscanthus*, creating an indirect competition with food crops. Streamflow is greatly reduced even at the watershed outlet and the effects of dry years are exacerbated (Stage 2, Figure 5.4). Ensuring minimum flow to urban water supply reservoirs results in worsening low flow conditions at the outlet (Stage 3, Figure 5.4) due to the sub-optimal spatial distribution of *Miscanthus*; production is shifted to the Salt Creek tributary and the Lower Sangamon. These are all trade-offs that would need to be taken into account to find the balance between the water needs of both the human and natural systems.

For the Kaskaskia watershed, historical impacts and underlying geology combine to reduce

the amount of NO_3-N exported here in spite of nearly 80% of the land devoted to annual row crops. Greater topographical relief provides more natural drainage, lessening the need for subsurface tile. The two large flood control reservoirs also have a major effect on NO_3-N removal in part because of their function to store spring flood flows, when a large portion of the annual NO_3-N export occurs in this region, for slow release during the summer. Thus in Kaskaskia a 45% NO_3-N reduction means an average of 7096 mT/yr less nutrient export and 37% of agricultural land converted to *Miscanthus*, while in Sangamon, the same constraint means an average of 11518 mT/yr less NO_3-N export and about half the watershed converted to fuel crops. Further analysis shows that the 30% reduction target means on average, 4730 mT/yr less NO_3-N export from Kaskaskia with 25% cropland conversion, while for Sangamon it means about 7678 mT/yr less nutrient export, on average and about 30% agricultural land converted. This amount of reduction is similar in magnitude to the amount of reduction Kaskaskia experiences at the 45% target, and for a similar percentage of land, which may explain in part why it appears that greater NO_3-N reduction can be achieved with less cropland conversion in Kaskaskia than in Sangamon.

Water quality goals could be a major driver of human decisions in Sangamon, but more care is needed in that watershed to ensure sufficient streamflow for both human and natural systems due to the location of water supply reservoirs relative to the optimal locations for growing *Miscanthus*. In Kaskaskia, however, it may be that low flows are a greater concern, especially at the outlet, due to minimum flow requirements for navigation there and on the Mississippi River. The Lake Carlyle reservoir might serve to mitigate these low flow effects in the driest years, as it has done in the recent severe drought, but this may not be a good long-term solution. Furthermore, the flood control reservoirs in Kaskaskia already provide a NO_3-N reduction benefit to that watershed, and since most of the optimal locations for *Miscanthus* do no affect flow into the reservoirs, the water quality benefit provided by fuel crops is in addition to that from the reservoirs. Thus the absolute reduction for Kaskaskia is much less than that for Sangamon. Within IL, watersheds with large reservoirs tend to export much less NO_3-N overall than those without (David et al., 2006). This suggests the possibility of coordination between watersheds – a low-export watershed pairing with a high-export one, the NO_3-N reduction target applied to the sum of the average annual

BAU export, and each watershed take half of the resulting value as the constraint, thus easing the burden on the high-export watershed, and taking advantage of reservoir-assisted denitrification.

While the results presented here do highlight the importance of location and history both within and between watersheds, especially the effects of water supply vs. flood control reservoirs, the results for Kaskaskia must be taken with some caution. Even if we assume the watershed being modeled is an extensively row-cropped version of the actual watershed, which would explain some of the excess NO_3-N export, some hillslope processes may still be poorly represented due to calibration with the forced single land use per subwatershed assumption. Both the comparative data and modeling analyses have highlighted the importance of heterogeneity in both surface and subsurface features of Kaskaskia in determining the catchment response to external inputs, and which, therefore, must be captured in the model representation of this watershed in order to reproduce these responses. However, the suboptimal hydrological representation of the Kaskaskia watershed does not explain why meeting the ethanol production mandate results in >30% reduction in NO_3-N export. Here, cellulosic ethanol rather than corn ethanol was more profitable, while corn for food/feed was more profitable than either type of ethanol, behavior which is the complete opposite of what was seen in the Sangamon watershed (Housh et al., 2014). Due to the subsystem interdependencies built into the SoS optimization model, the final optimized result can sometimes be unexpected, as illustrated by the results from Stage 3 of Trajectory 2 in Sangamon (Yaeger et al., 2014). Nonetheless, the model inputs and connections within and between the subsystems must be subject to data justification to ensure that this particular unexpected behavior is reasonable. That work is ongoing and beyond the scope of this thesis.

5.5 Conclusions

This study expanded on previous work, which developed an integrated modeling framework which coupled a human decisions model with a watershed model in an interactive way that was able to capture the co-evolutionary dynamics of the coupled human-hydrologic system as it spiraled forward in time to produce likely trajectories of future change. This

framework was then applied to the comparative analysis of the propagation of different biofuel-related external drivers through the human-hydrologic system in one intensively-managed Midwestern US agricultural watershed. In this paper, these same external drivers were imposed on a second, nearby, watershed with a different human-hydrologic system with a different co-evolutionary history. Again, because of the connection between agricultural development in the study region and the problem of hypoxia in the Gulf of Mexico as reflected in the region's known past historical trajectory, the complexity of human system responses to the imposed biofuel mandate was constrained by a secondary driver consisting of an annual watershed NO_3-N reduction target. Imposition of the external biofuel mandate combined with different secondary environmental targets on each watershed resulted in nonlinear and multi-scale responses of both the human and hydrologic systems, thus altering the future co-evolution of the system in complex, heterogeneous ways.

Results showed that both location and history are important when dealing with these coupled systems since the co-evolution of humans with the natural system results in unique patterns of change that influence the internal responses to imposition of external drivers. In this region at least, small (watershed) scale studies are necessary to study these interactions, because each watershed will have a unique mosaic of change, even if they are geographically close and climatically similar. This was illustrated by the very different effects the two types of reservoirs had on the internal responses of the coupled system to imposition of the same primary and secondary external drivers, at least for the largest NO_3-N reduction. Confidence in further conclusions from these results are hampered by the model representations of both systems; once these are satisfactorily improved, the nonlinearities in both the human and hydrologic responses, as well as the differences in trade-offs between the two watersheds will become more clear.

Acknowledgments

The research presented here was done as part of the NSF-funded project "Interdependence, Resilience, and Sustainability of Infrastructures for Biofuel Development" (NSF grant EFRI-083598, X. Cai, PI). We also thank Mark B. David for providing the upstream water quality

data used to calibrate the SWAT model in this study.

Chapter 6

Conclusions

The primary goal of this dissertation has been knowledge discovery, specifically, to develop predictive insights into the dynamic nature of the interactions between the two components of a coupled human and environmental system at the watershed scale through comparative analysis of historical data and within an integrated modeling framework.

6.1 Major Findings

The major findings of this dissertation can be organized around two aspects which together form the combined approach to the research: insights into catchment response to change from comparative data analysis of watershed historical data, and insights into human-hydrologic system response to change from comparative modeling analysis.

First, a large-scale comparative analysis of the flow regimes and flow duration curves in nearly 200 catchments across the Continental US was undertaken. The visual mapping done between the results from empirical, classification, and modeling aspects of the overarching study showed that within each regime class, both process and empirical, there was a great deal of heterogeneity in the shapes of the FDC which was most noticeable in the tails, which contain information about extreme flows. While the upper tail is most affected by the precipitation intensity, the lower tail is affected more by catchment properties of landscape, soils, geology, etc than by climate. Thus further analysis of the tails of the FDC, especially the low flows, may separate the physical controls on the shape of the FDC more effectively.

This understanding was applied to a comparative study of current data and historical behavior between two adjacent, intensively-managed, Midwestern US agricultural watersheds: Sangamon to the north, and Kaskaskia to the south. Despite their similarity in location,

drainage area, and climate, significant hydrologic differences were found both between the watersheds and within them. These differences resulted from the co-evolution over time of the human and hydrologic systems, through human adaptations to the climate, landscape, and hydrology. Analyses of hydrologic signatures revealed three main controls on the hydrologic response of these two watersheds. First, spatio-temporal variability of precipitation increased the heterogeneity of the catchment response. Second, inline reservoirs affected the catchment response differently, depending on their purpose: flood control reservoirs reduced the flood peaks and increased low flows of downstream reaches, and removed a large portion of NO_3-N inputs due to long residence times. Water supply reservoirs with short residence times reduced low flows somewhat, but had little effect on water quality. Third, at the scale of this study, tile drainage was associated with a higher BI and increased heterogeneity in the low flows, relative to less-tiled areas. However, because tile drainage in this region is also generally associated with intensive row-crop agriculture, it may be the combined effect of these land modifications that was observed. Analysis of the histories of the case study catchments showed that modifications and impacts have been layered on top of each other through time: fire, prairie conversion, surface drainage, subsurface (tile) drainage, reservoirs, intensification of row cropping, erosion, fertilizer application, etc., which resulted in the formation of a different spatial mosaic of change in each watershed. Thus even though a similar climatic external forcing is imposed on both watersheds, the internal responses of each watershed are different. Together, these differences would also affect the suitability of certain locations for growing Miscanthus, and thus must be taken into account in future plans for biofuels expansion to avoid worsening or creating water stress conditions.

The second aspect of the research involved development of a new quantitative framework for analysis of the dynamic interactions within coupled human and nature systems, and its application, in the context of biofuel development, to both the Sangamon and Kaskaskia watersheds. This framework combined a System of Systems optimization model for human decisions with respect to biofuel development with a watershed model for hydrologic and water quality feedbacks in an interactive way, capturing the co-evolutionary dynamics of the coupled system as it spiraled forward in time and produced likely trajectories of future change. Because of the connection between agricultural development in the study region and

the problem of hypoxia in the Gulf of Mexico as reflected in the regions known past historical trajectory, the complexity of human system responses to the imposed biofuel mandate was constrained by a secondary driver consisting of an annual watershed NO3-N reduction target. Imposition of the external biofuel mandate combined with different secondary environmental targets resulted in nonlinear and multi-scale responses of both the human and hydrologic systems, including multiple trade-offs, and impacted the future co-evolution of the system in complex, heterogeneous ways.

Although multiple of future trajectories can be generated with this integrated modeling tool, depending on human adaptations and changes to human values and norms, here only two such trajectories were explored. Both trajectories were developed progressively in three decadal stages, allowing for variable human feedbacks in response to the biofuel and nitrate outcomes produced in the preceding stage. In addition, the same cellulosic ethanol mandate was imposed on both trajectories, but a different initial secondary NO3-N reduction target was imposed on each: 45% for Trajectory 1 and 30% for Trajectory 2. In both cases the beneficial, targeted reduction in NO3-N also came with significant detrimental reduction in land area for food crops, annual profits, and summer low flows, all of which were smaller in scope in the latter case. For Trajectory 1, the detrimental reductions were noticeable at the outlet (the watershed scale), in both Sangamon and Kaskaskia, and prompted a subsequent adaptive response to these feedbacks at the same scale, with the humans reducing the NO_3-N target at the outlet to only 20% to balance out the benefits and detrimental effects. In the case of Trajectory 2, however, the control shifted to water deficits at the local (subwatershed) scale with respect to city water supplies in the Sangamon watershed, but remained at the outlet for the Kaskaskia watershed, since there were no water supply reservoirs in this watershed and the downstream reaches required sufficient flow for navigation. The two examples show that the spatial mosaic of human modifications to the hydrologic system is the first order control on the internal responses of the coupled human-hydrologic system to the imposed biofuel mandate and nitrate constraints. For the Sangamon watershed, the magnitude of the external drivers and constraints changes how they are manifested in the resulting dynamic changes in land use, average annual profits, and both water quality and quantity, thus showing the reciprocal feedbacks between human and natural systems. The

Kaskaskia watershed was less sensitive to the magnitude of the secondary constraint, since there was less NO_3-N export overall to reduce, due to the mitigating effects of large flood control reservoirs and less tile drainage within the watershed.

The trajectories explored in this analysis showed that the external drivers were propagated through the coupled system in complex ways at different scales, manifesting differently in different places. Because both model components of the adopted framework were distributed, the heterogeneity of both the human and hydrological systems was able to be captured. The interdependencies between the various subsystems in the SoS model created outcomes for the optimal biofuels land use development which then propagated through the hydrologic model from upstream to downstream, creating further complexity in the response. Furthermore, the coupled co-evolutionary modeling framework was designed to mimic the progressive human system responses to the hydrologic feedbacks, thus allowing for the possibility of course corrections along the way. Because of this, the process of developing the trajectories provided valuable predictive insights into the dynamics of the coupled system. For the case study system, the result of using the cellulosic ethanol mandate to reduce NO_3-N export to the Gulf was not a simple trade-off of water quantity for water quality, but instead involved multiple trade-offs affecting multiple subsystems and stakeholders, notably the trade-off between land for food crops and land for fuel crops, and for Sangamon, additional trade-offs between different water users within the watershed, such as cities, power plants, and discharge permit holders. The richness of the results possible from the use of the adopted framework also reflects the interaction between the external drivers and the internal mosaic of previous human modifications to the landscape and hydrology over the watersheds entire co-evolutionary history.

6.2 Limitations and Future Research

While the results presented here do highlight the importance of location and history both within and between watersheds, especially the effects of water supply vs. flood control reservoirs, the results for Kaskaskia must be taken with some caution. The poor calibration results indicated that a disconnect between the observed watershed and the modeled repre-

sentation, and thus some hydrologic processes were not being correctly represented. Both the comparative data and modeling analyses have highlighted the importance of heterogeneity in both surface and subsurface features of Kaskaskia in determining the catchment response to external inputs, and which, therefore, must be captured in the model representation of this watershed in order to reproduce these responses. However, the SoS model representation of Kaskaskia also produced some results that at seemed counterintuitive. Therefore the model inputs and connections within and between the subsystems must be verified and tested to determine if this unexpected behavior is reasonable. That work is ongoing and beyond the scope of this thesis.

The system trajectories chosen to demonstrate the application of this integrated modeling framework are only two examples of the many possible ways this system could co-evolve, depending on shifts in climate, development decisions, and response to hydrologic feedbacks. From the analysis done for the first part of this dissertation, it was found that in the past, this region has experienced periodic severe droughts, some extending for several years. Further exploration using this integrated model framework would impose drought conditions of various lengths and severity along with the RFS2 mandate, constrained by secondary environmental targets, to analyze the effects on the co-evolution of hydrologic processes and watershed management practices. Although the agriculture and biofuel development studied here were rainfed crops, an extended drought period would exacerbate the detrimental low flow effects seen under current climate conditions. Under water stress conditions, not only would environmental responses (crop yields, streamflow) be affected, socioeconomic priorities influencing the outcomes of the various trade-offs may change as well. Analysis of system co-evolution trajectories under these conditions could reveal additional insights into the coupled system behavior.

This study mainly considered human water use needs, with ecological needs assumed to be met with sufficient minimum flow. If the needs of environmental users such as specific fish functional guilds and other aquatic ecosystems are considered as well as the known human water demands, additional locations within the watershed may require sufficient flow or flow characteristics, which would further change the spatial configuration of biofuels development. Thus, in addition to the competition and trade-offs already discussed, there exists, but is

not considered in this study, a trade-off between ecological and human water users. While the ecological side has been ignored in the past, leading to some ecological degradation, human system values have been changing. Under the 1972 Clean Water Act, the USEPA has been overseeing ecosystem monitoring and restoration, and ecosystem integrity is now being accounted for in watershed management institutions. Thus the human system has shown some historical response to ecological feedbacks as well as purely hydrological ones. The modeling framework should be adapted to take account of these interactions as well, since they will likely further influence the future co-evolution of the coupled system.

How to incorporate these ecological needs into the hydrological system forms the third main theme of this research, How to use ecogeomorphology based on some simple geomorphologic parameters to translate streamflow from a hydrologic model into ecological habitat indicators?

In keeping with the goal of process-based understanding, we use geomorphologic data and fish sampling data to look for a correlation between hydraulic geometry and fish ecological indices in order to incorporate both ecology and hydrology in the quantitative assessment of the state of the environmental system in the framework.

Appendix A

Towards a more holistic assessment of the state of the natural system: more than minimum flow requirements

A.1 Introduction

Within the integrated modeling framework developed in this study, the states of both the human and environmental systems are assessed at the end of each decadal stage in order to determine impacts on and resulting feedbacks from each system. The state of the human system can be determined in economic terms, e.g. system profit. There are various methods for assessment of the state of the environmental system (Tharme, 2003), ranging from purely hydrological, using elements of the flow regime (Poff et al., 1997; Richter et al., 1997; Yang et al., 2008), to geomorphological, incorporating physical habitat measurements (Karim et al., 1995; Jowett, 1997; Brierley et al., 2010), or ecological, using relationships derived from biotic data (Hawkins et al., 2010; Bailey et al., 2012). Ideally, assessment of ecosystem health would involve some combination of hydrologic variability, habitat characteristics, and ecological information (Schaeffer et al., 1988; Niu et al., 2012), but some of this data can be difficult to collect and is often very site-specific (Karim et al., 1995).

Perhaps the simplest to implement, and consequently most widely used, method is to specify a minimum flow requirement, since low flow periods, especially those in summer, can be a time stress to aquatic life (Schlosser, 1991). While this method has an advantage as it is easily incorporated into the optimization model as a flow constraint, it also has limitations, since the presence of flow does not guarantee sufficient habitat, and in ephemeral conditions, pools can provide refuge habitat (Imhol et al., 1996; Welcomme et al., 2006). In the context of the biofuels case study, conversion of agricultural land to growing dedicated biofuel crops reduced streamflows, with summer low flow periods reduced to no-flow periods. While the smallest reaches were most affected, they also may not contain much aquatic life;

however, when water supplies to the human system were affected, large scale ecosystem damage may have already occurred. Between these two extremes, damage may occur that needs to be addressed but might be missed with only a simple flow threshold. An indicator that incorporates ecology as well as hydrology aid in identifying these ecosystem feedbacks, thus providing a more complete picture of the state of the environmental system.

To demonstrate the integrated modeling framework in this study, the state of the natural system is determined in three ways: by examining the flow duration curves (FDCs), the 7 day minimum flow (7-day minQ), and a simple flow deficit analysis which utilizes a minimum threshold flow, the water yield (mm) that over the time period of interest is exceeded 85% of the time (Q_{85}). All three be obtained from the output of the watershed hydrological model and were chosen for their relevance to low flows. In addition, two of these, FDCs and the 7-day minQ, are also used in the Indicators of Hydrologic Alteration (IHA) analysis of rivers, which relates statistical hydrologic parameters to ecosystem impacts (Richter et al., 1997). For the present study, to keep with the goal of understanding how these systems interact, we propose a simple process-based approach, using geomorphological properties.

In many cases, physical attributes of the stream reach are collected along with fish survey data (e.g., Table 1 in Cao and Hawkins (2005)). A comprehensive set of empirical relations have been developed for rivers in the state of IL, including the two case study rivers, that allow hydraulic geometry properties to be calculated from flow discharge (Stall and Fok, 1968); these have been applied to flow studies in these and nearby watersheds (Bhowmik, 1979; Larson et al., 1994; White et al., 2003). In addition, the SWAT model can be modified to output the simplified modeled hydraulic geometry, which can then be directly used to assess stream habitat, as has been done in the Sangamon river (Singh and Broeren, 1989). To connect the fish to the habitat, fish indices based on functional guilds (Welcomme et al., 2006) which group fish species by how they interact with their environment are used. Each guild would prefer or be able to adapt to a different kind of habitat and thus could be used to determine the ecological condition of the river.

The goal of the last part of the research is to connect stream ecology (fish) indicators to hydrology/geomorphology parameters to better describe the condition of the environmental system and will address two hypotheses:

1. Large scale planting of *Miscanthus* in these watersheds will further reduce summer low flows, and modeling results so far do show this. By determining environmental response using some minimum flow threshold and computing deficit volume, deficit period duration, and time between deficits assumes that by ensuring some minimum flow, all water users' (both human and environmental) needs will be satisfied. However, during low flow periods and droughts streamflow amount may not be as important to aquatic life such as fish as flow depth, or presence of pools ("refugia").

2. Hydraulic geometry combined with habitat characteristics may be better indicators of ecological 'health' than just minimum flow, or at least provide a more complete picture of the state of the environmental system.

Thus hydraulic geometry could provide the link between the watershed hydrologic model and stream ecology. This can be tested using regression analysis with habitat and hydraulic geometry parameters as predictors and several fish indices as response variables.

A.2 Methods

Stream ecology sampling data

Fish sampling is usually done during the low flow period and some habitat information is gathered at the same time (Holtrop et al., 2010). Of two datasets, fish indicators and habitat characteristics, provided for the two case study watersheds, only locations where there was data for both datasets were used. Thus when missing data points for habitat characteristics were removed, only 25 locations in Sangamon and 42 locations for Kaskaskia had a complete set of x and y variables. In addition, the data at each location for most of the habitat characteristics represented one point in space at one point in time. Spatial patterns of select predictors are presented in Figure A.1. Note that a large value of IBI is "good", while a large value for %TS is "bad" (a higher proportion of tolerant species can indicate that a given reach is in poor ecological condition, since species in this assemblage are better adapted to poor conditions than other species that might be normally found there). The majority

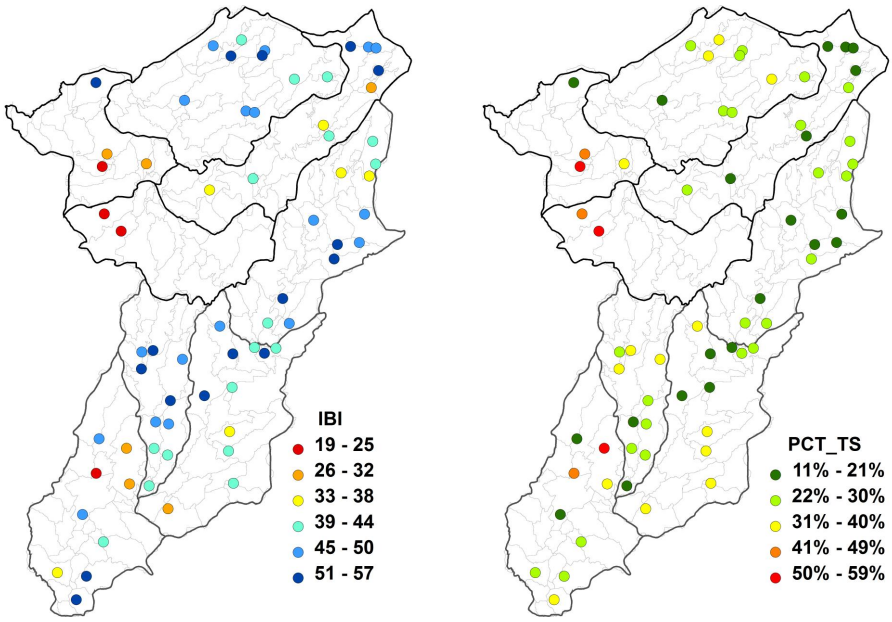

Figure A.1: Index of Biotic Integrity (left) and % Tolerant Species (right) for both watersheds.

locations in both watersheds have low proportions of tolerant species and high IBI, which is good. Lower Sangamon, Salt Fork Sangamon, and some parts of Lower Kaskaskia appear to be "problem" areas according to these indicators.

The final variables chosen for the analysis are listed in Table 1, grouped by type. For the fish indicators, these included IBI and component indicators of IBI. For the hydrology, these included hydraulic parameters (hydrologic model outputs), reach characteristics, some of which may be obtainable from land use and DEM datasets, and bed substrate characteristics, which might be roughly estimated from soils data sets.

A.2.1 Multivariable linear regression

A variable selection scheme, outlined in Ye et al. (2013) was followed to eliminate redundant or insignificant predictors. Because of the very small sample size, a small portion of the data set was used to test the prediction; i.e. 80% of the data was randomly selected to train the regression model and the remaining 20% was used to validate the prediction. This

Table A.1: Data used in the regression analysis

Response Variables (Y)		**Predictor Variables (X)**		
NF_SR	native Fish species richness	flow WIDTH(m)	can all be	Hydraulic variables
NM_SR	native minnow species richness	flow Velocity(m/s)	obtained	
NS_SR	native sucker species richness	flow DEPTH(m)	from SWAT	
NSF_SR	native sunfish species richness	PCT_CANOPY	reach riparian canopy cover	Reach variables
NBI_SR	native benthic invertevore species richness	PCT_POOL	refugia; low V, large D	
NI_SR	native intolerant species richness	PCT_RIFFLE	high V, low D	
		PCT_RUN	between pool and riffle	
PCT_SBIS	proportion of specialist benthic invertevore species	PCT_MUD		Bed substrate variables
		PCT_SAND		
PCT_GF	proportion of generalist feeders	PCT_FGRAVEL	fine gravel	
		PCT_MGRAVEL	medium gravel	
P_OCMSS	proportion of obligate-coarse-mineral-substrate spawning species	PCT_CGRAVEL	coarse gravel	
		PCT_SCOBBLE	small cobbles	
PCT_TS	proportion of tolerant species	PCT_LCOBBLE	large cobbles	
		PCT_BEDROCK		
IBI	Index of Biotic Integrity	PCT_CLAYPAN		

process was repeated 50 times, with R^2 calculated each time, and an average R^2 calculated for all runs. The best fit indicators in each watershed obtained from this analysis were then compared to determine which variables were important in each watershed, and if they were the same variables.

Three statistical criteria were used to select the variables: Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and the least absolute shrinkage and selection operator (LASSO). The final predictor variables in the regression relationship were those where the frequency each predictor was chosen was greater than 50% for all three methods, and of these, those with >90% frequency selection in the LASSO method. In addition, in order to link hydrologic model output to these response variables, one of the predictor variables needed to be one of the hydraulic geometry variables (flow width, depth, or velocity). The watersheds were analyzed separately, since previous work (Yaeger et al., 2013) has highlighted differences in hydrologic response, as well as in ecological indicator patterns (Cao, pers. comm.) between the two watersheds.

A.3 Results

Before commencing multivariable regression, simple linear regression was performed to determine if there were any interesting patterns or relationships in the data (Figures A.2, A.3, A.4, and A.5).

Surprisingly, there did seem to be some possible relationships. For example, in scatter plots of the predictor variable "flow width" and all 11 response variables, there did appear to be some relationship with %TS, especially in Kaskaskia (Figure A.2). For flow width,this is perhaps less surprising that there was a relationship, since hydraulic geometry tends to vary predictably through the stream network, with reaches becoming deeper and wider as stream order and drainage area increases. In general, xy scatter plots showed more structure in Kaskaskia except for the width-depth (WD) ratio; here Sangamon scatterplots appeared more structured (Figure A.3), since the Sangamon WD-ratios covered a much smaller range of values than Kaskaskia.

In general, the results of the multivariate linear regression showed that in both watersheds,

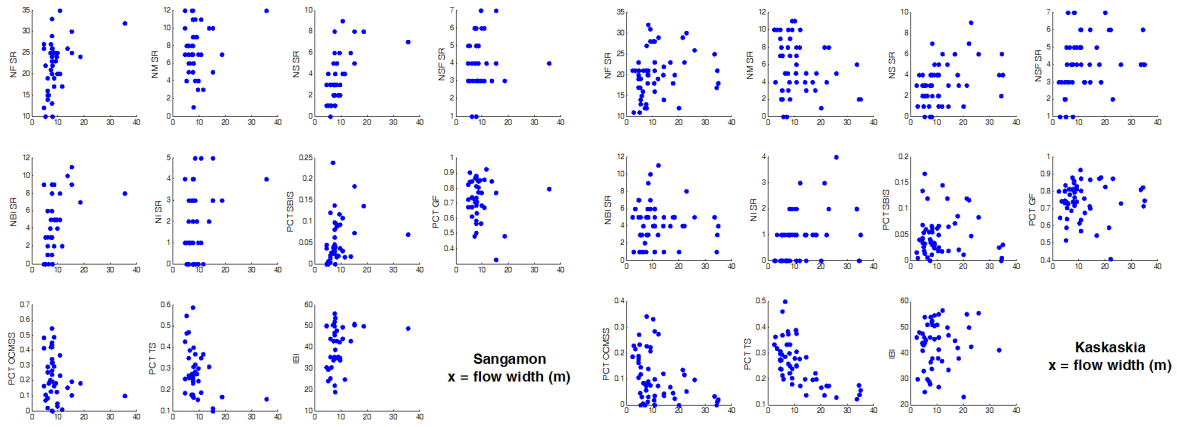

Figure A.2: Scatter plots of the predictor variable flow width and all 11 response variables for Sangamon (left) and Kaskaskia (right).

NF, NM, NS, and NBI species richness were all predicted fairly well (R^2 >0.5, p <0.001). Additionally, in Sangamon, both %TS and IBI, the overall fish ecological indicator, also were predicted well; in Kaskaskia, only IBI. In Kaskaskia, flow velocity (m/s) is important in predicting IBI, but in Sangamon, no hydraulic variables are important. Thus there is no possible connection with SWAT model output in Sangamon for IBI.

For the Sangamon watershed, the best index with a connection to hydraulic geometry was Native Sucker Species Richness (count):

$$y = -1.008 + 0.467x_1 - 2.16x_2 + 10.18x_3 \tag{A.1}$$

where y = NS SR, x_1 = flow width (m), x_2 = fraction of riparian canopy, and x_3 = % fine gravel substrate. Here, flow width is the main connection to the hydrologic model. Other parameters can be interpolated spatially from data points. Alternatively, bed substrate could be estimated from soils data, % canopy could be estimated from land use/land cover data set (possibly). A scatterplot of the predicted native sucker species richness vs. observed richness is shown in Figure A.6.

For the Kaskaskia watershed the best index was the Native Benthic Invertivore Species Richness (count):

$$y = 2.58 + 0.113x_1 + 19.49x_2 - 5.59x_3 - 5.12x_4 \tag{A.2}$$

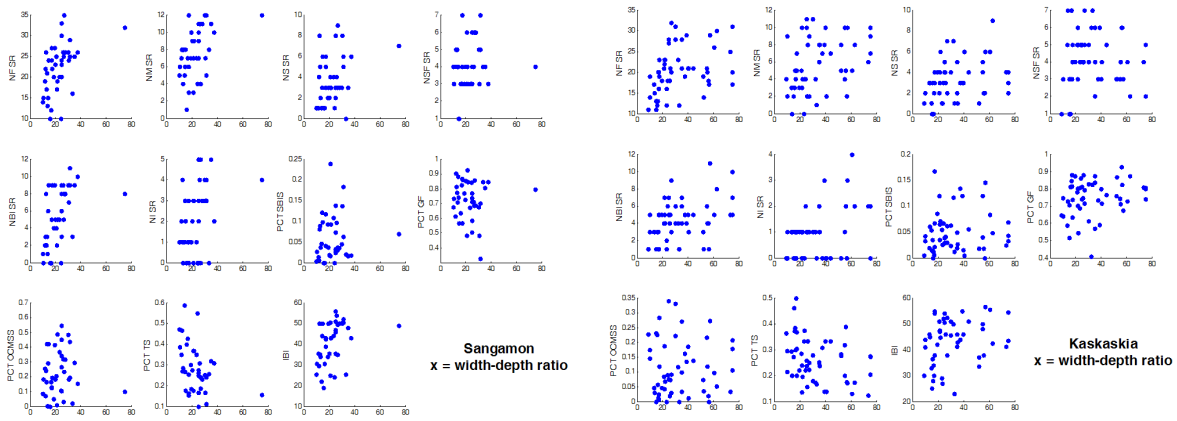

Figure A.3: Scatter plots of the predictor variable WD-ratio and all 11 response variables for Sangamon (left) and Kaskaskia (right).

where y = NBI SR and x_1 = flow width (m), x_2 = % riffle in the reach, x_3 = % medium gravel substrate, and x_4 = % claypan substrate. Flow width is also the main connection to the hydrologic model, as it was in Sangamon. %Riffle may be important here due to the higher slopes and more varied topography in Kaskaskia compared to Sangamon. A scatterplot of the predicted native sucker species richness vs. observed richness is shown in Figure A.7. With more data points in its dataset, Kaskaskia also had good fits with IBI, %Tolerant Species and others.

The number of native sucker species was <5 in most reaches in both watersheds, but Sangamon had a greater proportion of reaches with the highest numbers (8 - 11) of benthic feeding species than Kaskaskia. These high values were in the upper reaches of Sangamon and Salt Creek (north), but were distributed further south in Kaskaskia (Figure A.8).

A.4 Discussion

An interesting result of this analysis was that not only are different indicators better in each watershed, but the importance of predictors was also different between watersheds (Figure A.9). For example, the hydraulic variables, which could be a link between a hydrologic model and the ecologic data, was not an important predictor in all cases; some response variables,

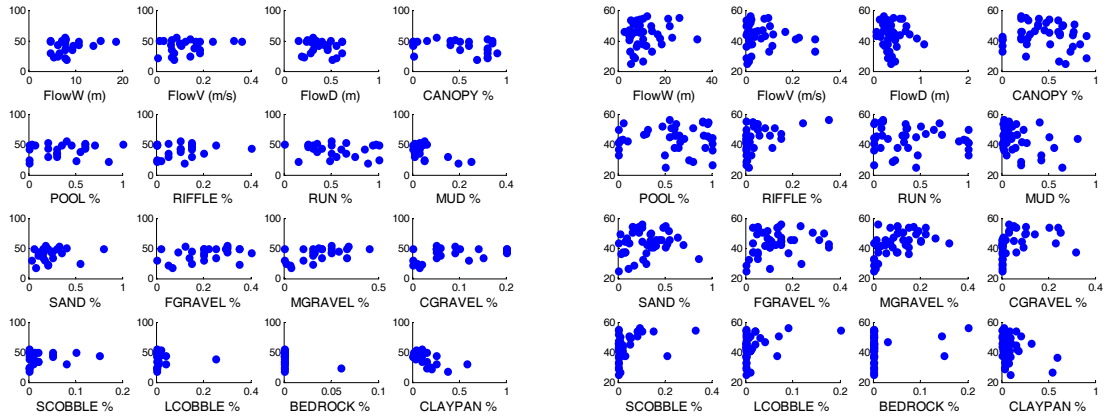

Figure A.4: Scatter plots of the response variable IBI and all predictor variables for Sangamon (left) and Kaskaskia (right).

while showing a strong relationship, did not depend on flow width, depth, or velocity. The proportion of the reach estimated to be pools was not considered an important predictor of any of the 11 ecological indicators, for either watershed. This may be due to the streams in these watersheds, especially Sangamon, being highly channelized, with a trapezoidal cross section in many cases (Frothingham et al., 2002). Bed substrate size was important for both watersheds for most indicators; overall, in Sangamon, mud, bedrock and claypan (finer particle sizes) were more important while in Kaskaskia it was gravels and cobbles (larger particle sizes).

Of the possible predictors chosen here, bed substrate characteristics were chosen most often, followed by hydraulic variables. Other habitat characteristics such as amount of vegetation and woody detritus present in the reach were not chosen for this analysis since they could not be estimated from land cover data. The importance of bed substrate seen here is similar to the results found for suburban watersheds where a more sophisticated regression analysis showed that while IBI did not correlate with urbanization land use change parameters, but rather with bed substrate size and channel slope (Fitzpatrick et al., 2005).

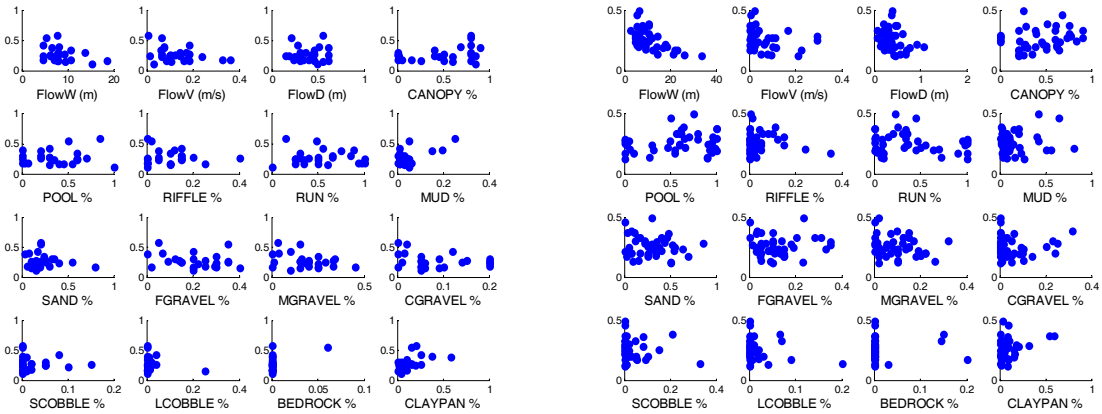

Figure A.5: Scatter plots of the response variable %tolerant species and all predictor variables for Sangamon (left) and Kaskaskia (right).

A.5 Application

From the modeled trajectories of change, land use change is predicted to decrease summer low flows, thus reducing flow width and depth. The results of the regression analysis provide an equation that can be used to predict an ecological indicator using habitat characteristics and hydraulic geometry, where the hydrologic model will provide daily hydraulic geometry. There are some questions regarding limitations of this applying these relationships to the trajectories of change, however, that need to be addressed first.

One limitation is that NS ranges from 0 to 11 in the data and represents fish counts (integer), and especially in Kaskaskia, different values of X (habitat data) produce same Y value (for most indicators). Thus other factors besides the hydraulic geometry are involved. Also, the fish response may not occur immediately; there is a cross-timescale lag between the human impact and the ecological feedback response. Furthermore, the hydrologic model produces daily flow width; in light of the time lag, this is not very meaningful. Thus some representative value is needed, such as the summer minimum, an average over the low flow period, etc. Lastly, as was mentioned previously, the data the regression on which the regression was based represented one point in space at one point in time. Thus any variation in hydraulic geometry is in space only, within and between the watersheds. For example, flow width tends to be smaller in headwater reaches, increasing downstream as drainage area

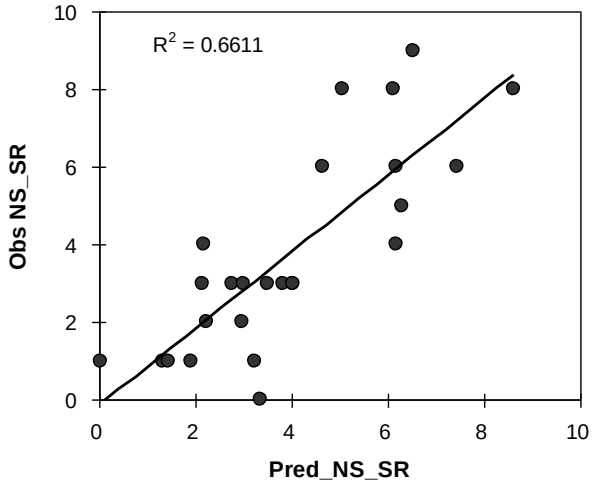

Figure A.6: Scatterplot of obs vs best-predicted indicator Sangamon, Native Sucker Species Richness.

increases, and this pattern of variation is explained by natural network processes, not land use change. Variation in ecological indices in this case could be due to different fish species living in upstream compared to downstream. This relationship could be thought of as a space for time substitution, in which case reduced flow would make larger reaches more like upstream ones, but the trajectories showed that large streams were less affected than small ones.

A.6 Conclusions

The linear relationships developed here are statistically significant and fairly strong, but are they physically meaningful/useful? More sophisticated statistical analysis can be tried, but time is limiting factor. This analysis could be greatly improved with a data set that varies in time as well as space, to understand better the historical response of fish species to the changes already taking place.

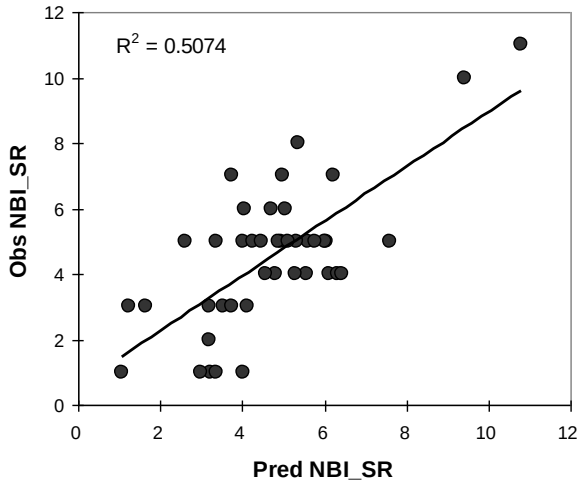

Figure A.7: Scatterplot of obs vs best-predicted indicator for Kaskaskia, Native Benthic Invertivore Species Richness.

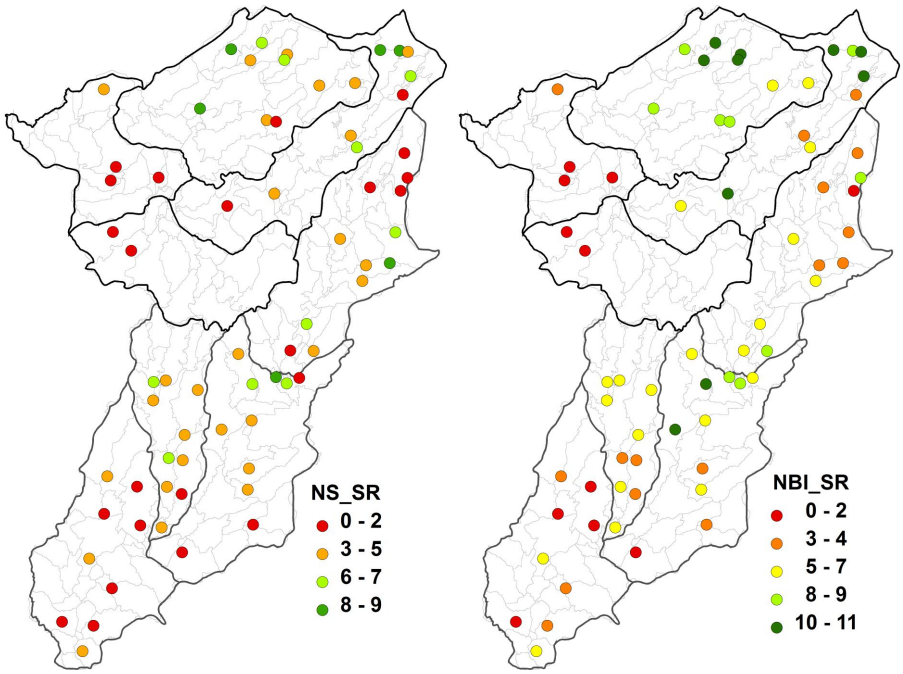

Figure A.8: Spatial distribution of the best-fit predictors for both watersheds; NBI (left) and NS (right).

Response Variables (Y)	Predictor Variables (X)																
	Hydraulic Variables			Reach and Refuge				Bed Substrate Variables									
	flow W (m)	flow V (m/s)	flow D (m)	% Canopy	% Pool	% Riffle	% RUN	% MUD	% SAND	% Fine Gravel	% Med. Gravel	% Coarse Gravel	% Small Cobble	% Large Cobble	% Bed rock	% Clay pan
NF_SR	K			S			K	K		K			K		S	S
			S					S			S	S				K
NM_SR				K						K			K		S	
NS_SR	S	S		S						S						
	K					K				K	K					
NSF_SR						S								S		S
NBI_SR								S				S				S
	K	K				K		K			K	K		K		K
NI_SR	K															
PCT_SBIS	S		S						S							S
						K										
PCT_GF	S								K							
P_OCMSS			S					S			S					
											K					
PCT_TS								S							S	S
	K															K
IBI								S			S				S	S
		K				K			K		K		K			

	NF_SR	NM_SR	NS_SR	NSF_SR	NBI_SR	NI_SR	%SBIS	%GF	% OCMSS	%TS	IBI	
r^2	**0.576**	**0.576**	**0.688**	0.471	**0.563**	---	0.394	0.192	0.452	**0.596**	**0.561**	Sangamon
p-value	0.0000	0.0000	0.0000	0.0002	0.0000	---	0.0008	0.0283	0.0002	0.0000	0.0000	
r2	**0.584**	**0.507**	**0.508**	---	**0.602**	0.336	0.125	0.135	0.134	0.428	**0.543**	Kaskaskia
p-value	0.0000	0.0000	0.0000	---	0.0000	0.0001	0.0216	0.0166	0.0170	0.0000	0.0000	

Figure A.9: Comparison between watersheds of predictor variable importance for different response variables

References

Arnold, J. and Fohrer, N. (2005). SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes*, 19(3):563–572.

Arora, M., Goel, N., Singh, P., and Singh, R. (2005). Regional flow duration curve for a Himalayan river Chenab. *Nordic Hydrology*, 36:193–206.

Bailey, R. C., Scrimgeour, G., Coté, D., Kehler, D., Linke, S., and Cao, Y. (2012). Bioassessment of stream ecosystems enduring a decade of simulated degradation: lessons for the real world. *Canadian Journal of Fisheries and Aquatic Sciences*, 69(4):784–796.

Baron, J. S., Poff, N. L., Angermeier, P. L., Dahm, C. N., Gleick, P. H., Hairston Jr, N. G., Jackson, R. B., Johnston, C. A., Richter, B. D., and Steinman, A. D. (2002). Meeting ecological and societal needs for freshwater. *Ecological Applications*, 12(5):1247–1260.

Basu, N. B., Destouni, G., Jawitz, J. W., Thompson, S. E., Loukinova, N. V., Darracq, A., Zanardo, S., Yaeger, M., Sivapalan, M., Rinaldo, A., and Rao, P. S. C. (2010). Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophysical Research Letters*, 37(23):L23404, doi:10.1029/2010GL045168.

Basu, N. B., Thompson, S. E., and Rao, P. S. C. (2011). Hydrologic and biogeochemical functioning of intensively managed catchments: A synthesis of top-down analyses. *Water Resources Research*, 47(10):W00J15, doi:10.1029/2011WR010800.

Baumhardt, R. L. (2003). The Dust Bowl era. In Stewart, B. A. and Howell, T. A., editors, *Encyclopedia of Water Science*, pages 187–191. Marcel-Dekker, NY.

Bhowmik, N. G. (1979). *Hydraulics of flow in the Kaskaskia River, Illinois*. Illinois State Water Survey.

Bogue, A. (1994). *From prairie to corn belt: farming on the Illinois and Iowa prairies in the nineteenth century*. Iowa State University Press, reprint edition.

Bogue, M. B. (1951). The Swamp Land Act and wetland utilization in Illinois, 1850-1890. *Agricultural History*, 25(4):169–180.

Borah, D., Bera, M., and Shaw, S. (2003). Water, sediment, nutrient, and pesticide measurements in an agricultural watershed in Illinois during storm events. *Transactions of the ASAE*, 46(3):657–674.

Botter, G., Porporato, A., Daly, E., Rodriguez-Iturbe, I., and Rinaldo, A. (2007a). Probabilistic characterization of base flows in river basins: Roles of soil, vegetation, and geomorphology. *Water Resources Research*, 43(6):W06404, doi:10.1029/2006WR005397.

Botter, G., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2007b). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. *Water Resources Research*, 43(2):W02417, doi:10.1029/2006WR005043.

Botter, G., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2009). Nonlinear storage-discharge relations and catchment streamflow regimes. *Water Resources Research*, 45(10):W10427, doi:10.1029/2008WR007658.

Botter, G., Zanardo, S., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2008). Ecohydrological model of flow duration curves and annual minima. *Water Resources Research*, 44(8):W08418, doi:10.1029/2008WR006814.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1993). *Classification and Regression Trees*. Chapman and Hall, Boca Raton.

Brierley, G., Reid, H., Fryirs, K., and Trahan, N. (2010). What are we monitoring and why? Using geomorphic principles to frame eco-hydrological assessments of river condition. *Science of the Total Environment*, 408(9):2025–2033.

Brooks, P. D., Troch, P. A., Durcik, M., Gallo, E., and Schlegel, M. (2011). Quantifying regional scale ecosystem response to changes in precipitation: Not all rain is created equal. *Water Resources Research*, 47(10):W00J08 doi: 10.1029/2010WR009762.

Cai, X., McKinney, D. C., and Lasdon, L. S. (2002). A framework for sustainability analysis in water resources management and application to the Syr Darya Basin. *Water Resources Research*, 38(6):21–1 – 21–14.

Cai, X., McKinney, D. C., and Lasdon, L. S. (2003). Integrated hydrologic-agronomic-economic model for river basin management. *Journal of Water Resources Planning and Management*, 129(1):4–17.

Cao, Y. and Hawkins, C. P. (2005). Simulating biological impairment to evaluate the accuracy of ecological indicators. *Journal of Applied Ecology*, 42(5):954–965.

Carpenter, S. R., Stanley, E. H., and Vander Zanden, M. J. (2011). State of the world's freshwater ecosystems: physical, chemical, and biological changes. *Annual Review of Environment and Resources*, 36:75–99.

Carrillo, G., Troch, P., Sivapalan, M., Wagener, T., Harman, C., and Sawicz, K. (2011). Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient. *Hydrology and Earth System Sciences*, 15(3):3411–3430.

Castellarin, A., Camorani, G., and Brath, A. (2007). Predicting annual and long-term flow-duration curves in ungauged basins. *Advances in Water Resources*, 30(4):937–953.

Castellarin, A., Vogel, R. M., and Brath, A. (2004). A stochastic index flow model of flow duration curves. *Water Resources Research*, 40(3):W03104, doi:10.1029/2003WR002524.

Cheng, L., Yaeger, M., Viglione, A., Coopersmith, E., Ye, S., and Sivapalan, M. (2012). Exploring the physical controls of regional patterns of flow duration curves – Part 1: Insights from statistical analyses. *Hydrology and Earth System Sciences*, 16(6):4435–4446.

Chiang, L.-C., Yuan, Y., Mehaffey, M., Jackson, M., and Chaubey, I. (2014). Assessing SWAT's performance in the Kaskaskia River watershed as influenced by the number of calibration stations used. *Hydrological Processes*, 28(3):676–687.

Chiu, Y.-W., Walseth, B., and Suh, S. (2009). Water embodied in bioethanol in the United States. *Environmental Science & Technology*, 43(8):2688–2692.

Claps, P. and Fiorentino, M. (1997). Probabilistic flow duration curves for use in environmental planning and management. In Harmancioglu, N. B., Alpaslan, M. N., Ozkul, S. D., and Singh, V. P., editors, *Integrated Approach to Environmental Data Management Systems*, volume 2 of *NATO ASI Series*, pages 255–266. Springer, Kluwer Academy.

Collins, S. L., Carpenter, S. R., Swinton, S. M., Orenstein, D. E., Childers, D. L., Gragson, T. L., Grimm, N. B., Grove, J. M., Harlan, S. L., Kaye, J. P., Knapp, A. K., Kofinas, G. P., Magnuson, J. J., McDowell, W. H., Melack, J. M., Ogden, L. A., Robertson, G. P., Smith, M. D., and Whitmer, A. C. (2010). An integrated conceptual framework for long-term social-ecological research. *Frontiers in Ecology and the Environment*, 9(6):351–357.

Connell, D. and Grafton, R. Q. (2011). Water reform in the Murray-Darling Basin. *Water Resources Research*, 47(12):WR009820, doi:10.1029/2010WR009820.

Cook, B. I., Miller, R. L., and Seager, R. (2009). Amplification of the North American Dust Bowl drought through human-induced land degradation. *Proceedings of the National Academy of Sciences*, 106(13):4997–5001.

Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L., and Sivapalan, M. (2012). Exploring the physical controls of regional patterns of flow duration curves – Part 3: A catchment classification system based on seasonality and runoff regime. *Hydrology and Earth System Sciences*, 16:4467–4482.

David, B., McIsaac Gregory, F., Royer Todd, V., Darmody Robert, G., and Gentry Lowell, E. (2001). Estimated historical and current nitrogen balances for Illinois. *The Scientific World Journal*, 1:597–604.

David, M. B., Drinkwater, L. E., and McIsaac, G. F. (2010). Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, 39(5):1657–1667.

David, M. B., Flint, C. G., McIsaac, G. F., Gentry, L. E., Dolan, M. K., and Czapar, G. F. (2013). Biophysical and social barriers restrict water quality improvements in the Mississippi River Basin. *Environmental Science & Technology*, 47(21):11928–11929.

David, M. B., Wall, L. G., Royer, T. V., and Tank, J. L. (2006). Denitrification and the nitrogen budget of a reservoir in an agricultural landscape. *Ecological Applications*, 16(6):2177–2190.

De Fraiture, C., Giordano, M., and Liao, Y. (2008). Biofuels and implications for agricultural water use: blue impacts of green energy. *Water Policy*, 10:67–81.

DeFries, R. and Eshleman, K. (2004). Land-use change and hydrologic processes: a major focus for the future. *Hydrological Processes*, 18(11):2183–2186.

DeFries, R. S., Foley, J. A., and Asner, G. P. (2004). Land-use choices: balancing human needs and ecosystem function. *Frontiers in Ecology and the Environment*, 2(5):249–257.

Demissie, Y., Yan, E., and Wu, M. (2012). Assessing regional hydrology and water quality implications of large-scale biofuel feedstock production in the Upper Mississippi river basin. *Environmental Science & Technology*, 46(16):9174–9182.

Dohleman, F. G., Heaton, E. A., and Long, S. P. (2010). Perennial grasses as second-generation sustainable feedstocks without conflict with food production. In *Handbook of Bioenergy Economics and Policy*, pages 27–37. Springer.

Donner, S. D. and Kucharik, C. J. (2008). Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. *Proceedings of the National Academy of Sciences*, 105(11):4513–4518.

Donner, S. D., Kucharik, C. J., and Foley, J. A. (2004). Impact of changing land use practices on nitrate export by the Mississippi River. *Global Biogeochemical Cycles*, 18(1):GB1028, doi:10.1029/2003GB002093.

Donner, S. D. and Scavia, D. (2007). How climate controls the flux of nitrogen by the Mississippi River and the development of hypoxia in the Gulf of Mexico. *Limnology and Oceanography*, 52(2):856–861.

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets, F., Hall, A., Hay, L., et al. (2006). Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, 320(1):3–17.

Egbendewe-Mondzozo, A., Swinton, S. M., Izaurralde, C. R., Manowitz, D. H., and Zhang, X. (2011). Biomass supply from alternative cellulosic crops and crop residues: A spatially explicit bioeconomic modeling approach. *Biomass and Bioenergy*, 35(11):4636–4647.

Falkenmark, M. (1997). Society's interaction with the water cycle: a conceptual framework for a more holistic approach. *Hydrological Sciences Journal*, 42(4):451–466.

Falkenmark, M. (2003). Freshwater as shared between society and ecosystems: from divided approaches to integrated challenges. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1440):2037–2049.

Feng, X., Vico, G., and Porporato, A. (2012). On the effects of seasonality on soil water balance and plant growth. *Water Resources Research*, 48(5):W05543, doi:10.1029/2011WR011263.

Fitzpatrick, F. A., Diebel, M. W., Harris, M. A., Arnold, T. L., Lutz, M. A., and Richards, K. D. (2005). Effects of urbanization on the geomorphology, habitat, hydrology, and fish index of biotic integrity of streams in the Chicago area, Illinois and Wisconsin. In *American Fisheries Society Symposium*, volume 47, pages 87–115.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K. (2005). Global consequences of land use. *Science*, 309(5734):570–574.

Frothingham, K. M., Rhoads, B. L., and Herricks, E. E. (2002). A multiscale conceptual framework for integrated ecogeomorphological research to support stream naturalization in the agricultural Midwest. *Environmental Management*, 29(1):16–33.

Ganora, D., Claps, P., Laio, F., and Viglione, A. (2009). An approach to estimate nonparametric flow duration curves in ungauged basins. *Water Resources Research*, 45(10):W10418, doi:10.1029/2008WR007472.

Gassman, P., Reyes, M., Green, C., and Arnold, J. (2007). The soil and water assessment tool: Historical development, applications, and future research directions. *Transactions of the ASABE*, 50(4):1211–1250.

Gentry, L. E., David, M. B., Below, F. E., Royer, T. V., and McIsaac, G. F. (2009). Nitrogen mass balance of a tile-drained agricultural watershed in east-central Illinois. *Journal of Environmental Quality*, 38(5):1841–1847.

Gerbens-Leenes, W., Hoekstra, A. Y., and van der Meer, T. H. (2009). The water footprint of bioenergy. *Proceedings of the National Academy of Sciences*, 106(25):10219–10223.

Goolsby, D. A., Battaglin, W. A., Lawrence, G. B., Artz, R. S., Aulenbach, B. T., Hooper, R. P., Keeney, D. R., and Stensland, G. J. (1999). Flux and sources of nutrients in the Mississippi-Atchafalaya River Basin. In *Topic 3 report for the integrated assessment on hypoxia in the Gulf of Mexico*, NOAA Coastal Ocean Program Analysis Series, page 130. NOAA Coastal Ocean Office, Silver Springs, MD.

Gordon, L. J., Peterson, G. D., and Bennett, E. M. (2008). Agricultural modifications of hydrological flows create ecological surprises. *Trends in Ecology & Evolution*, 23(4):211–219.

Goswami, D., Kalita, P., Cooke, R., and Hirschi, M. (2008). Estimation and analysis of baseflow in drainage channels in two tile-drained watersheds in Illinois. *Transactions of the ASABE*, 51(4):1201–1213.

Harman, C. and Troch, P. (2014). What makes Darwinian hydrology" Darwinian"? Asking a different kind of question about landscapes. *Hydrology and Earth System Sciences*, 18(2):417–433.

Hawkins, C. P., Olson, J. R., and Hill, R. A. (2010). The reference condition: predicting benchmarks for ecological and water-quality assessments. *Journal of the North American Benthological Society*, 29(1):312–343.

Heaton, E. A., Dohleman, F. G., and Long, S. P. (2008). Meeting US biofuel goals with less land: the potential of Miscanthus. *Global Change Biology*, 14(9):2000–2014.

Hickman, G. C., Vanloocke, A., Dohleman, F. G., and Bernacchi, C. J. (2010). A comparison of canopy evapotranspiration for maize and two perennial grasses identified as potential bioenergy crops. *GCB Bioenergy*, 2(4):157–168.

Hisdal, H. and Tallaksen, L. M. (2003). Estimation of regional meteorological and hydrological drought characteristics: a case study for Denmark. *Journal of Hydrology*, 281(3):230–247.

Holmes, M., Young, A., Gustard, A., Grew, R., et al. (2002). A region of influence approach to predicting flow duration curves within ungauged catchments. *Hydrology and Earth System Sciences*, 6(4):721–731.

Holtrop, A. M., Cao, Y., and Dolan, C. R. (2010). Estimating sampling effort required for characterizing species richness and site-to-site similarity in fish assemblage surveys of wadeable Illinois streams. *Transactions of the American Fisheries Society*, 139(5):1421–1435.

Housh, M., Cai, X., Ng, T. L., McIsaac, G. F., Ouyang, Y., Khanna, M., Sivapalan, M., Jain, A., Eckhoff, S., Gasteyer, S., Al-Qadi, I., Bai, Y., Yaeger, M. A., Ma, S., and Song, Y. (2013). System of systems model for biofuel development analysis. *Journal of Infrastructure Systems*. In review.

Housh, M., Yaeger, M., Cai, X., McIsaac, G. F., Khanna, M., Sivapalan, M., Jain, A., Ouyang, Y., and Al-Qadi, I. (2014). Managing multiple mandates: a System of Systems model to analyze strategies for producing cellulosic ethanol and reducing riverine nitrate loads in the upper Mississippi River Basin. *submitted to Environmental Science & Technology*.

Hu, X., McIsaac, G., David, M., and Louwers, C. (2007). Modeling riverine nitrate export from an east-central Illinois watershed using SWAT. *Journal of Environmental Quality*, 36(4):996–1005.

Imhol, J., Fitzgibbon, J., and Annable, W. (1996). A hierarchical evaluation system for characterizing watershed ecosystems for fish habitat. *Canadian Journal of Fisheries and Aquatic Sciences*, 53(S1):312–326.

Jackson, L. (2002). Restoring prairie processes to farmlands. In Jackson, D. L. and Jackson, L. L., editors, *The Farm as Natural Habitat*, pages 137–154. Island Press, Washington, D.C.

Jain, A. K., Khanna, M., Erickson, M., and Huang, H. (2010). An integrated biogeochemical and economic analysis of bioenergy crops in the Midwestern United States. *GCB Bioenergy*, 2(5):217–234.

Jowett, I. (1997). Instream flow methods: a comparison of approaches. *Regulated Rivers: Research & Management*, 13(2):115–127.

Kalita, P., Algoazany, A., Mitchell, J., Cooke, R., and Hirschi, M. (2006). Subsurface water quality from a flat tile-drained watershed in Illinois, USA. *Agriculture, Ecosystems & Environment*, 115(1):183–193.

Karim, K., Gubbels, M. E., and Goulter, I. (1995). Review of determination of instream flow requirements with special application to Australia. *Water Resources Bulletin*, 31(6):1063–1077.

Karlen, D. L., Dinnes, D. L., and Singer, J. W. (2010). Midwest soil and water conservation: Past, present, and future. In Zobeck, T. M. and Schillinger, W. F., editors, *Soil and Water Conservation Advances in the US: Past Effects Future Outlook*, pages 131–162. Soil Science Society of America, Inc., Madison, WI.

Keefer, L., Bauer, E., and Markus, M. (2010). Hydrologic and nutrient monitoring of the Lake Decatur Watershed: Final report 19932008. Contract Report 2010-07, Illinois State Water Survey, Champaign, IL.

King, C. W., Holman, A. S., and Webber, M. E. (2008). Thirst for energy. *Nature Geoscience*, 1(5):283–286.

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, 42(3):W03S04, doi:10.1029/2005WR004362.

Klemeš, V. (1988). A hydrological perspective. *Journal of Hydrology*, 100(1):3–28.

Larson, R. S., Butts, T. A., and Singh, K. P. (1994). *Water Quality and Habitat Suitability Assessment: Sangamon River between Decatur and Petersburg*. Illinois State Water Survey.

Le, P. V., Kumar, P., and Drewry, D. T. (2011). Implications for the hydrologic cycle under climate change due to the expansion of bioenergy crops in the Midwestern United States. *Proceedings of the National Academy of Sciences*, 108(37):15085–15090.

LeBoutillier, D. W. and Waylen, P. R. (1993). A stochastic model of flow duration curves. *Water Resources Research*, 29(10):3535–3541.

Levy, S. (2003). Turbulence in the Klamath River basin. *BioScience*, 53(4):315–320.

Li, H., Sivapalan, M., Tian, F., and Liu, D. (2010a). Water and nutrient balances in a large tile-drained agricultural catchment: a distributed modeling study. *Hydrology and Earth System Sciences*, 14(3):2259–2275.

Li, M., Shao, Q., Zhang, L., and Chiew, F. H. (2010b). A new regionalization approach and its application to predict flow duration curve in ungauged basins. *Journal of Hydrology*, 389(1):137–145.

Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., Pell, A. N., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C. L., Schneider, S. H., and Taylor, W. (2007a). Complexity of coupled human and natural systems. *Science*, 317(5844):1513–1516.

Liu, J., Dietz, T., Carpenter, S. R., Folke, C., Alberti, M., Redman, C. L., Schneider, S. H., Ostrom, E., Pell, A. N., Lubchenco, J., Taylor, W. W., Ouyang, Z., Deadman, P., Kratz, T., and Provencher, W. (2007b). Coupled human and natural systems. *AMBIO: A Journal of the Human Environment*, 36(8):639–649.

L'vovich, M. I. (1979). *World Water Resources and Their Future*. American Geophysical Union, Washington, D.C. English translation by R. L. Nace.

Martin, J. (2011). Perspective: Don't foul the water. *Nature*, 474(7352):S17–S17.

McIsaac, G. and Hu, X. (2004). Net N input and riverine N export from Illinois agricultural watersheds with and without extensive tile drainage. *Biogeochemistry*, 70(2):253–273.

McIsaac, G. F., David, M. B., and Mitchell, C. A. (2010). Miscanthus and switchgrass production in central Illinois: Impacts on hydrology and inorganic nitrogen leaching. *Journal of Environmental Quality*, 39(5):1790–1799.

Milly, P., Betancourt, J., Kalnenmark, M., Hirsch, R., Kundewicz, Z., Lettenmaier, D., and Stouffer, R. (2008). Stationarity is dead: Whither water management? *Science*, 319(5863):573–574.

Moraes, M. M. G. A., Cai, X., Ringler, C., Albuquerque, B. E., Vieira da Rocha, S. P., and Amorim, C. A. (2009). Joint water quantity-quality management in a biofuel production area Integrated economic-hydrologic modeling analysis. *Journal of Water Resources Planning and Management*, 136(4):502–511.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., and Veith, T. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3):885–900.

Muneepeerakul, R., Azaele, S., Botter, G., Rinaldo, A., and Rodriguez-Iturbe, I. (2010). Daily streamflow analysis based on a two-scaled gamma pulse model. *Water Resources Research*, 46(11):W11546, doi:10.1029/2010WR009286.

Nash, J. and Sutcliffe, J. (1970). River flow forecasting through conceptual models: part I. a discussion of principles. *Journal of Hydrology*, 10(3):282–290.

Neitsch, S., Arnold, J., Kiniry, J., Srinivasan, R., and Williams, J. (2010). *Soil and water assessment tool: Input/output file documentation, Version 2009*. USDA-ARS Grassland, Soil, and Water Research Laboratory, Temple, TX.

Ng, T. L., Eheart, J. W., Cai, X., and Miguez, F. (2010). Modeling Miscanthus in the Soil and Water Assessment Tool (SWAT) to simulate its water quality effects as a bioenergy crop. *Environmental Science & Technology*, 44(18):7138–7144.

Niu, S. Q., Franczyk, M. P., and Knouft, J. H. (2012). Regional species richness, hydrological characteristics and the local species richness of assemblages of North American stream fishes. *Freshwater Biology*, 57(11):2367–2377.

Panno, S. V., Kelly, W. R., Hackley, K. C., Hwang, H.-H., and Martinsek, A. T. (2008). Sources and fate of nitrate in the Illinois River Basin, Illinois. *Journal of Hydrology*, 359(1):174–188.

Parker, D. C., Hessl, A., and Davis, S. C. (2008). Complexity, land-use modeling, and the human dimension: Fundamental challenges for mapping unknown outcome spaces. *Geoforum*, 39(2):789–804.

Peel, M. C. and Blöschl, G. (2011). Hydrological modelling in a changing world. *Progress in Physical Geography*, 35(2):249–261.

Perrone, D. and Hornberger, G. M. (2014). Water, food, and energy security: scrambling for resources or solutions? *Wiley Interdisciplinary Reviews: Water*, 1(1):49–68.

Phillips, J. D. (2001). Human impacts on the environment: unpredictability and the primacy of place. *Physical Geography*, 22(4):321–332.

Phillips, J. D. (2002). Global and local factors in earth surface systems. *Ecological Modelling*, 149(3):257–272.

Phillips, J. D. (2003). Sources of nonlinearity and complexity in geomorphic systems. *Progress in Physical Geography*, 27(1):1–23.

Phillips, J. D. (2006). Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change. *Hydrology & Earth System Sciences*, 10(5):731–742.

Phillips, J. D. (2007). The perfect landscape. *Geomorphology*, 84(3):159–169.

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E., and Stromberg, J. C. (1997). The natural flow regime. *BioScience*, 47(11):769–784.

Prince, H. (2008). *Wetlands of the American Midwest: a historical geography of changing attitudes*. University of Chicago Press, Chicago.

Raymond, P. A., David, M. B., and Saiers, J. E. (2012). The impact of fertilization and hydrology on nitrate fluxes from Mississippi watersheds. *Current Opinion in Environmental Sustainability*, 4(2):212–218.

Richter, B., Baumgartner, J., Wigington, R., and Braun, D. (1997). How much water does a river need? *Freshwater Biology*, 37(1):231–249.

Rockström, J., Falkenmark, M., Karlberg, L., Hoff, H., Rost, S., and Gerten, D. (2009). Future water availability for global food production: the potential of green water for increasing resilience to global change. *Water Resources Research*, 45(7):W00A12, doi:10.1029/2007WR006767.

Royer, T. V., David, M. B., and Gentry, L. E. (2006). Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: Implications for reducing nutrient loading to the Mississippi River. *Environmental Science & Technology*, 40(13):4126–4131.

Royer, T. V., Tank, J. L., and David, M. B. (2004). Transport and fate of nitrate in headwater agricultural streams in Illinois. *Journal of Environmental Quality*, 33(4):1296–1304.

Sauquet, E. and Catalogne, C. (2011). Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in france. *Hydrology and Earth System Sciences*, 15(2):24212435.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P., and Carrillo, G. (2011). Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences*, 15(3):2895–2911.

Scavia, D. and Donnelly, K. A. (2007). Reassessing hypoxia forecasts for the Gulf of Mexico. *Environmental Science & Technology*, 41(23):8111–8117.

Schaeffer, D. J., Herricks, E. E., and Kerster, H. W. (1988). Ecosystem health: I. measuring ecosystem health. *Environmental Management*, 12(4):445–455.

Scheffer, M., Westley, F., and Brock, W. (2003). Slow response of societies to new problems: causes and costs. *Ecosystems*, 6(5):493–502.

Schilling, K. E., Jha, M. K., Zhang, Y.-K., Gassman, P. W., and Wolter, C. F. (2008). Impact of land use and land cover change on the water balance of a large agricultural watershed: Historical effects and future directions. *Water Resources Research*, 44(7):2007WR006644, doi:10.1029/2007WR006644.

Schlosser, I. J. (1991). Stream fish ecology: a landscape perspective. *BioScience*, 41(10):704–712.

Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., and Bacmeister, J. T. (2004). On the cause of the 1930s Dust Bowl. *Science*, 303(5665):1855–1859.

Secchi, S., Gassman, P. W., Jha, M., Kurkalova, L., and Kling, C. L. (2011). Potential water quality changes due to corn expansion in the Upper Mississippi River Basin. *Ecological Applications*, 21(4):1068–1084.

Singh, K. P. and Broeren, S. M. (1989). Hydraulic geometry of streams and stream habitat assessment. *Journal of Water Resources Planning and Management*, 115(5):583–597.

Sivapalan, M. (2006). Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. In Anderson, M. G. and McDonnell, J. J., editors, *Encyclopedia of Hydrological Sciences*. John Wiley and Sons, Inc.

Sivapalan, M., Savenije, H. H., and Blöschl, G. (2012). Socio-hydrology: A new science of people and water. *Hydrological Processes*, 26(8):1270–1276.

Sivapalan, M., Thompson, S., Harman, C., Basu, N., and Kumar, P. (2011a). Water cycle dynamics in a changing environment: Improving predictability through synthesis. *Water Resources Research*, 47(10):2011WR011377, doi:10.1029/2011WR011377.

Sivapalan, M., Yaeger, M. A., Harman, C. J., Xu, X., and Troch, P. A. (2011b). Functional model of water balance variability at the catchment scale: 1. evidence of hydrologic similarity and space-time symmetry. *Water Resources Research*, 47(2):2010WR009568,doi:10.1029/2010WR009568.

Smith, C. M., David, M. B., Mitchell, C. A., Masters, M. D., Anderson-Teixeira, K. J., Bernacchi, C. J., and DeLucia, E. H. (2013). Reduced nitrogen losses after conversion of row crop agriculture to perennial biofuel crops. *Journal of Environmental Quality*, 42(1):219–228.

Smith, R. A., Alexander, R. B., and Lanfear, K. J. (1990). Stream water quality in the conterminous United States Status and trends of selected indicators during the 1980s. *National Water Summary*, 91:111–140.

Sprague, L. A., Hirsch, R. M., and Aulenbach, B. T. (2011). Nitrate in the Mississippi River and its tributaries, 1980 to 2008: Are we making progress? *Environmental Science & Technology*, 45(17):7209–7216.

Stall, J. B. and Fok, Y. S. (1968). Hydraulic geometry of Illinois streams. Research Report 15, University of Illinois Water Resources Center.

Stemler, J. (2011). Water control operations. Presentation to the Kaskaskia Basin Water Supply Plan Committee, accessed 12/2013.

Tharme, R. E. (2003). A global perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers. *River Research and Applications*, 19(5-6):397–441.

Thomas, M. F. (2001). Landscape sensitivity in time and space – an introduction. *Catena*, 42(2):83–98.

Thompson, S., Sivapalan, M., Harman, C., Srinivasan, V., Hipsey, M., Reed, P., Montanari, A., and Blöschl, G. (2013). Developing predictive insight into changing water systems: use-inspired hydrologic science for the Anthropocene. *Hydrology & Earth System Sciences*, 17(12):5013–5039.

Trimble, S. W. (1985). Perspectives on the history of soil erosion control in the eastern United States. *Agricultural History*, 59(2):162–180.

Troch, P. A., Martinez, G. F., Pauwels, V., Durcik, M., Sivapalan, M., Harman, C., Brooks, P. D., Gupta, H., and Huxman, T. (2009). Climate and vegetation water use efficiency at catchment scales. *Hydrological Processes*, 23(16):2409–2414.

Turner, R. E. and Rabalais, N. N. (2003). Linking landscape and water quality in the Mississippi River Basin for 200 years. *BioScience*, 53(6):563–572.

van der Velde, Y., Rozemeijer, J. C., de Rooij, G. H., van Geer, F. C., and Broers, H. P. (2010). Field-scale measurements for separation of catchment discharge into flow route contributions. *Vadose Zone Journal*, 9(1):25–35.

Van Lanen, H., Wanders, N., Tallaksen, L., and Van Loon, A. (2013). Hydrological drought across the world: impact of climate and physical catchment structure. *Hydrology and Earth System Sciences*, 17(5):1715–1732.

Van Loon, A. and Van Lanen, H. (2013). Making the distinction between water scarcity and drought using an observation-modeling framework. *Water Resources Research*, 49:1483–1502.

Vanloocke, A., Bernacchi, C. J., and Twine, T. E. (2010). The impacts of Miscanthus x. giganteus production on the Midwest US hydrologic cycle. *GCB Bioenergy*, 2(4):180–191.

Varghese, S. (2007). Biofuels and global water challenges. *Institute for Agriculture and Trade Policy. Minnesota*, pages 1–7.

Vogel, R. M. and Fennessey, N. M. (1994). Flow-duration curves I: New interpretation and confidence intervals. *Journal of Water Resources Planning and Management*, 120(4):485–504.

Vogel, R. M. and Fennessey, N. M. (1995). Flow duration curves II: A review of applications in water resource planning. *JAWRA Journal of the American Water Resources Association*, 31(6):1029–1039.

Vörösmarty, C. J., McIntyre, P., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., and Liermann, C. (2010). Global threats to human water security and river biodiversity. *Nature*, 467(7315):555–561.

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V., Kumar, P., Rao, P. S. C., Basu, N. B., and Wilson, J. S. (2010). The future of hydrology: An evolving science for a changing world. *Water Resources Research*, 46(5):WR008906, doi:10.1029/2009WR008906.

Wallace, J. S., Acreman, M. C., and Sullivan, C. A. (2003). The sharing of water between society and ecosystems: from conflict to catchment–based co–management. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1440):2011–2026.

Walsh, R. and Lawler, D. (1981). Rainfall seasonality: description, spatial patterns and change through time. *Weather*, 36(7):201–208.

Wang, D. and Hejazi, M. (2011). Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States. *Water Resources Research*, 47(10):W00J12, doi:10.1029/2010WR010283.

Welcomme, R., Winemiller, K., and Cowx, I. (2006). Fish environmental guilds as a tool for assessment of ecological condition of rivers. *River Research and Applications*, 22(3):377–396.

White, A. B., Kumar, P., Saco, P. M., Rhoads, B. L., and Yen, B. C. (2003). Changes in hydrologic response due to stream network extension via land drainage activities. *JAWRA Journal of the American Water Resources Association*, 39(6):1547–1560.

Whitney, G. G. (1996). *From coastal wilderness to fruited plain: a history of environmental change in temperate North America from 1500 to the present*. Cambridge University Press, Cambridge.

Woodhouse, C. A. and Overpeck, J. T. (1998). 2000 years of drought variability in the central United States. *Bulletin of the American Meteorological Society*, 79(12):2693–2714.

Wooten, H. H. and Jones, L. A. (1955). The history of our drainage enterprises. In *Water, the yearbook of agriculture*, pages 478–491. U.S. Dept. of Agriculture, Washington, D.C.

Wu, M., Demissie, Y., and Yan, E. (2012). Simulated impact of future biofuel production on water quality and water cycle dynamics in the Upper Mississippi river basin. *Biomass and Bioenergy*, 41:44–56.

Wu, Y. and Liu, S. (2012). Impacts of biofuels production alternatives on water quantity and quality in the Iowa River Basin. *Biomass and Bioenergy*, 36:182–191.

Yadav, M., Wagener, T., and Gupta, H. (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, 30(8):1756–1774.

Yaeger, M., Coopersmith, E., Ye, S., Cheng, L., Viglione, A., and Sivapalan, M. (2012). Exploring the physical controls of regional patterns of flow duration curves – Part 4: A synthesis of empirical analysis, process modeling and catchment classification. *Hydrology and Earth System Sciences*, 16:4483–4498.

Yaeger, M., Sivapalan, M., McIsaac, G., and Cai, X. (2013). Comparative analysis of hydrologic signatures in two agricultural watersheds in east-central Illinois: legacies of the past to inform the future. *Hydrology and Earth System Sciences*, 17:4607–4623. doi:10.5194/hess-17-4607-2013.

Yaeger, M. A., Housh, M., Cai, X., and Sivapalan, M. (2014). Water for food, energy, and the environment: projecting the trajectory of flow regime and water quality with increasing biofuel crop production in the US Corn Belt. *submitted to Water Resources Research*.

Yang, Y.-C. E., Cai, X., and Herricks, E. E. (2008). Identification of hydrologic indicators related to fish diversity and abundance: A data mining approach for fish community analysis. *Water Resources Research*, 44(4):W04412, DOI: 10.1029/2006WR005764.

Ye, S., Li, H.and Huang, M., Ali, M., Leng, G., Leung, R. L.-Y., Wang, S., and Sivapalan, M. (2013). Subsurface stormflow parameterization for land surface models: derivation from regional analysis of streamflow recession curves. *Journal of Hydrology*. in review.

Ye, S., Yaeger, M., Coopersmith, E., Cheng, L., and Sivapalan, M. (2012). Exploring the physical controls of regional patterns of flow duration curves – Part 2: Role of seasonality, the regime curve, and associated process controls. *Hydrology and Earth System Sciences*, 16(11):4447–4465.

Yevjevich, V. (1967). *An objective approach to definitions and investigations of continental hydrologic droughts*. Colorado State University Fort Collins.

Yokoo, Y. and Sivapalan, M. (2011). Towards reconstruction of the flow duration curve: development of a conceptual framework with a physical basis. *Hydrology and Earth System Sciences*, 15(9):2805–2819.

Zelenhasić, E. and Salvai, A. (1987). A method of streamflow drought analysis. *Water Resources Research*, 23(1):156–168.

Zhang, Y.-K. and Schilling, K. (2006). Increasing streamflow and baseflow in Mississippi River since the 1940s: Effect of land use change. *Journal of Hydrology*, 324(1):412–422.